

onNote:

カメラ画像による紙楽譜認識を用いた演奏メディア

山本 祐介[†] 内山 英昭[‡] 笥 康明^{†††}

本研究では、カメラ画像による紙楽譜認識を用いた演奏メディアの提案を行う。演奏メディアとは、例えば、レコードや CD のように記録された音楽を用いて演奏できるメディアである。本研究では、紙の楽譜をそのまま演奏に用いる、直感的な演奏メディアを提案する。この演奏メディアでは、紙楽譜の楽譜としての視覚性、素材としての特性を活かし、カメラに対して楽譜を動かすことで記録された音楽を演奏することが出来る。具体的に本研究では、文章画像検索法 LLAH を応用し、カメラ画像からマーカレスの紙楽譜に対してその種類や姿勢を検出する楽譜画像検索法を提案し、それを用いて楽譜と音とを対応付け、演奏できるシステムを提案する。また、システムを用いた演奏方法、演奏に必要なビジュライジングの方法を検討する。今後、楽譜の折れや重なり認識方法などを検討することで、より直感的な演奏メディアとなり、様々な人の音楽創作活動を支援するプラットフォームとして価値のあるものになる。

onNote:

An Examination of Performance Media Using Paper Score Recognition with a Camera

YUSUKE YAMAMOTO[†] HIDEAKI UCHIYAMA[‡] YASUAKI KAKEHI^{†††}

This paper will discuss a type of performance media that uses paper score recognition with a camera. Performance media are media that perform music like that recorded on records and CDs, for example. The kind of intuitive performance media discussed here uses a musical score written on paper, visualizing the score based on the characteristics of the paper. Music can be performed by moving the score toward the camera. This paper will discuss use of a score image search method applying LLAH in which the camera detects the type and position of an unmarked score and performance is achieved by connecting sounds to the score. The method of using this system and that of achieving the visualization necessary for performance will also be examined. In the future, after developing ways to “read” paper scores that have fold marks and those partially obscured because they are on top of each other, performance media will become more intuitive and develop into a platform that can support the musical creativity of many artists.

1. はじめに

近年デジタルメディアを用いて、様々な楽器や音楽制作ツールが提案され、音楽を創る行為は、特別な教育を受けた人のみならず、より多くの人々が楽しめるものとなってきた。音楽をより身近なものにするために、レコードや CD に代表される演奏メディアは大きな役割を担っている。ここで、演奏メディアとは、予め記録された音楽を取り出して曲を構成する媒体を指す。これらは、ただ内包している音楽データを再生するのみならず、DJ が行うスクラッチなどのように簡易な

操作を加えることで、多様なエフェクトや表現につなげることが出来る。



図1 onNote

[†] 慶應義塾大学大学院 政策・メディア研究科
Graduate School of Media and Governance, Keio University

[‡] フランス国立情報学自動制御研究所
INRIA Rennes

^{†††} 慶應義塾大学環境情報学部
Faculty of Environment and information Studies, Keio University

しかし、レコード盤だけを見ても、その中のどこにどのような音が入っているのかが分からないように、これまでの演奏メディアの多くは記録された情報としての音楽と、それを記録する媒体となる物理的素材の間に明確な対応づけがなされておらず、直感的な操作という面で考えれば、十分なインタフェースと言えない。

このような背景のもと、本研究では、音が記号として記録された物理的媒体として楽譜（五線記譜法を用いた楽譜）に着目する。今回筆者らは紙の楽譜をそのまま演奏に用いる直感的な演奏メディア onNote を提案する（図 1）。ユーザは、カメラに楽譜をかざし、その前をスライドさせることで、自分の好きな速度で、好きな位置の音楽を演奏できる。楽譜は視覚的に音楽の存在や構造を読み取ることができる。また、楽譜が印刷された紙は私達の身近な素材であり、折る、曲げる、重ねるといった多様な使い方ができる。これらの操作を演奏手法として積極的に取り入れることで、高い直感性を有する演奏メディアの実現を目指す。

具体的に今回は、カメラ画像からマーカレスの紙楽譜に対してその種類や姿勢を検出する楽譜画像検索法とその検出方法を用いて楽譜を音と対応付け、演奏に必要な情報をビジュアルライジングするシステムを提案・実装した。本稿では、以下に関連研究をまとめた後に、システムの提案、実装、さらにはユーザの反応について述べる。

2. 関連研究

演奏のためのメディアは、普及しているレコードや CD の他に、インタフェース技術を用いて多様な提案がなされてきた。タンジブルなインタフェースを用いて演奏を行う代表例として、ブロックの位置や向きおよび組み合わせで音楽を作る *Reactable* [1] や *Small Fish* [2]、細切れにしたレコードを小型のロボットで走査して音楽を再生する *Sound Chaser* [3]、レーザー光とのインタラクションで音を奏でる *scoreLight* [4] などが挙げられる。このようなタンジブルな演奏メディアは演奏に古典楽器のような専門知識を必要とせず、記録された音楽を使って簡易な操作で音楽性の高い演奏をすることができるため、様々な人の音楽創作活動を支援したり、創作意欲を刺激する上で価値がある。

今回の我々の提案では、演奏操作をより直感的に行う上で、タンジブルインタフェースとして楽譜を記した紙を用いる。身近に手に入る楽譜をそのまま用いることで、電気的な加工やマーカなどの貼付を必要とし

ないこと、操作と出力される音との対応付けをユーザに明確に伝えられる点が特徴である。楽譜をポインティングすることにより演奏を行うメディアとしては、*Melody Assistant*[5]などが挙げられる。これらは、モニタ上に表示された楽譜を GUI で操作するものであり、折る、曲げる、重ねるといった物理的な紙ならではの操作を演奏に取り込み、直感性をあげる我々の研究とはアプローチが異なる。

紙楽譜の ID（楽曲の種類およびページ番号）を認識する方法は、従来からいくつかの取り組みがなされている[6]。それらの多くは、楽譜に含まれる演奏情報をデジタルで扱えるデータに変換することを目的としている。例えば、楽譜を MIDI 情報に変換するための技術である。変換の際のアルゴリズムとしては、音符や記号を抽出し、パターンマッチングを行うものである。現在は、精度と変換速度が向上し、市販のソフト[7]でも、1枚の楽譜を数秒でデジタル情報に変換できる。しかし、カメラ画像からの認識や紙が変形し隠れが生じた場合の認識などには不向きであり、今回のように楽譜をカメラの前で様々に動かしてその姿勢や ID をリアルタイムで読み取るような目的には適さない。そこで、本研究では楽譜の中に記された記号情報を用いて自然特徴点検索のアプローチで楽譜の姿勢や ID をリアルタイムに検出する手法を提案・実装する。

3. onNoteの提案

3.1 システム概要

本研究では、紙の楽譜を直接演奏に用い楽譜の音を紙の特性を用いてコントロールする。本システムでは、ユーザは紙の楽譜をカメラの前にかざし、動かすことにより、任意の場所の音楽を任意の速度で演奏できる。このような機能を実現するために、図 2 のようなシステムを提案する。

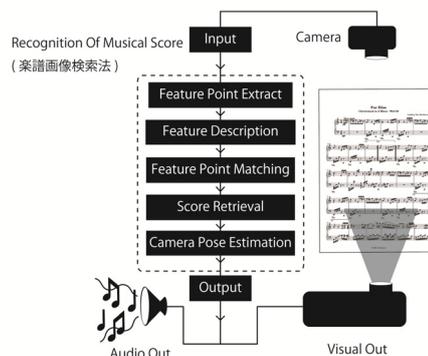


図2 システム基本構成

3.2 楽譜特徴点抽出処理

楽譜に限らず、マーカレスで物体の状態認識と特定を行う手段として、自然特徴点検索法が挙げられる。自然特徴点検索法とは、実世界に既に存在する特徴を用いたパターン認識の方法であり、人為的に画像処理のためのマーカを貼付する必要がある ARToolkit [8] のように対象物体に配置したマーカによって物体の表面情報を妨げる問題がない。また、得られる情報も多いため、カメラ画像からの検索が可能で、物体の射影変形や幾何学的に複雑な変形もカメラ画像から認識できるという利点もある。

本システムでは、自然特徴点検索法を用いて楽譜を認識することを考える。自然特徴点検索法として代表的なものとして、SIFT[9]や SURF [10] がある。しかし、これらの手法で用いられる特徴点、特徴量の計算はリッチなテクスチャに適した方法であり、楽譜のような二値画像に対して不向きである。二値画像の検索に適した手法としては、点や線などの幾何学位置関係を利用した GH(Geometric Hashing) [11] や、文章の検索法として用いられる LLAH [12] がある。この中で、GH は計算コストが非常に大きく、リアルタイムでの情報提示は難しい。一方、LLAH は文章に特化した特徴点抽出のために設計されており、楽譜をそのまま認識することは出来ないが、検索処理が高速であり、射影変形や照明変化に対してもロバストである。さらに文献[13]では、LLAH の検索アルゴリズムの高速化、特徴点のトラッキング手法が提案され、処理速度とカメラの位置に対する自由度が改善された。また、文献[14]では、検索画像の折れの検出が行われた。このような背景から、本研究では LLAH をベースに楽譜に適した特徴点抽出を行うことで認識を行う。以下に、本システムでの楽譜認識アルゴリズムを記す。

まず、LLAH は文章画像検索法であるため、楽譜で LLAH の特徴点抽出処理を行うと、記号が五線をまたいでいるため、段落が連結し、1つの楽譜から数個しか特徴点を得られない。自然特徴点検索法は特徴点を用いて検索を行うため、1つの画像が多くの特徴点を持っていたほうが望ましい。

そこで、楽譜の五線譜を含む細い線を削除し、楽譜を分割し、音符の旗や玉など分割された部分を特徴点とするアルゴリズムを導入する。具体的には、図3のように入力画像をグレースケールに変換、ガウシアンフィルタでぼかし、適性2値化で画像の濃く残った部分だけを抽出し、その重心を特徴点とする処理を施す。

提案システムでは、必要な情報をプロジェクタで直



図3 楽譜特徴点抽出処理

接楽譜に投影する。そのため、プロジェクタでの投影映像を単色とし、楽譜特徴点抽出処理の入力画像をグレースケールに変換する際に、その色を除き、認識の影響を避ける。

3.3 楽譜とMIDIの対応付け

前処理として、楽譜を本システムに登録し、印刷しておく、同時にデータベースの楽譜の小節ごとに MIDI を割り当てる。MIDI は小節の示すはじめの音の鳴り始めとなり終わりを区切りとして楽譜の xy 座標に割り当てる。演奏の際には、楽譜の再生したい座標を指定することで、その箇所に対応する MIDI が再生される。

4. onNoteにおける演奏手法

上記のような設計のもと、システムを実装した。ここでは、このシステムを用いた演奏手法を提案する。

4.1.1 演奏手法

演奏方法としては以下の3つの方法を提案する。

なお、楽譜に対するカメラの位置姿勢から、カメラの中心座標が楽譜のどこにあるのか計算できるため、これをポイント指定に用いる。

手法 1: 投影されたポイントに楽譜を合わせることで、その部分の音が鳴る。ユーザはポイントで楽譜をなぞるように演奏する。

手法 2: 楽譜のポイントを指定し、その部分から楽譜の指定した方向に指定した速度で再生し、改行、繰り返し記号、終了記号で終了する。

手法 3: 楽譜のポイントを2点指定し、2点間の音を指定した方向に指定した速度で再生する。



図4 演奏手法：手法1（左図）、手法2（中央）、手法3（右図）

提案する楽譜画像検索法は射影変換にロバストであるため、カメラを自由に動かすことができる。カメラと楽譜の位置関係によって、楽譜の紙の特性をどの程度反映できるかといったことが変わってくるため、それぞれにカメラを固定した方法と自由に動かせる方法が考えられる。

4.1.2 ビジュアルライジング

演奏に必要な情報としてポイント指定に用いる座標がある。本研究では、紙楽譜をインタフェースとして用いるため、プロジェクタを用いてポイントを楽譜に直接視覚化する。例えば、手法1の場合は、図5のように音がなった際に描かれている図形が変化し、ユーザが視覚的にフィードバックを得られるようにした。

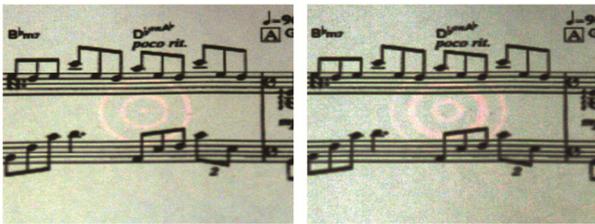


図5 投影されるポイント：音が鳴っていない状態（左図）、音が鳴っている状態（右図）

5. まとめ

本研究では、カメラ画像による紙楽譜認識を用いた演奏メディア onNote のシステム、紙楽譜の認識方法、提案した演奏メディアの演奏の提案を行った。

本研究で提案した onNote は、既存の楽譜という演奏の記号化形式を用いる。さらに、その楽譜が記された紙というタンジブルな素材の操作性を演奏方法として取り入れることで、高い直感性を有するメディアとなり、様々な人の音楽創作活動を支援するプラットフォームとして価値のあるメディアとなる。

今後は、まず提案した楽譜画像検索法と評価改良するなど基礎技術の向上に取り組む。さらに、基礎技術によって可能になる演奏方法を提案、実装評価する。また、現在のシステムはカメラに対する楽譜の回転などの姿勢の変化が演奏に反映できない、カメラの自由度に制限がある、楽譜の折れや重なりが検出できないなどの問題がある。そのため、文献[13]、文献[14]を参考にし、楽譜の折れや重なりを取り入れたより発展的なインタラクシオン技術を検討し、それらによって可能になる演奏方法を実装評価していく。例えば、楽譜の折れや重なりが検出できれば楽譜同士を折って、任意のシーケンスをつなげて演奏することもできるよ

うになり、さらに直感的な演奏が出来ると考える。さらに、現在手動で行っているデータベースの楽譜と MIDI との対応付けを、既存の楽譜認識を用いて自動化できるようにすることも今後の重要な課題である。

参考文献

- 1) Jorda, S., Geiger, G., Alonso, M. and Kaltenbrunner, M.: "The reacTable: Exploring the Synergy between Live Music Performance and Tabletop Interface", TEI'07, ACM Press, pp.139-146 (2007).
- 2) 藤幡正樹: "Small Fish", http://www.aec.at/bilderclient_detail_en.php?id=276, 2010年12月現在.
- 3) YURI SUZUKI: "Sound Chaser", <http://www.yurisuzuki.com/soundchaser.html>, 2010年12月現在.
- 4) Cassinell, A., Kuribara, Y., Zerroug, A., Manabe, D. and Ishikawa, M.: "scoreLight: playing with a human sized laser pickup", NIME2010, pp. 144-149 (2010).
- 5) Myriad Software: "Melody Assisrant", <http://www.myriad-online.com/en/products/melody.htm>, 2010年12月現在.
- 6) 加藤博一, 井口征士: "小節単位処理に基づいたピアノ楽譜の自動認識", 電子情報通信学会論文誌, vol.J71-D, No.5, pp.894-901(1988).
- 7) 株式会社河合楽器製作所: <http://www.kawai.co.jp/cm/music/products/sm/>, 2010年12月現在.
- 8) Kato, H. and Billinghurst, M.: "Marker Tracking and HMD Calibration for a video-based Augmented Reality Conferencing System", Proc IWAR'99, pp. 85-94 (1999).
- 9) Low, D.G.: "Distinctive image feature from scale invariant key points", IJCV, vol.60, pp91-110, 2004.
- 10) Bay, H., Ess, A., Tuytelaars, T. and Gool, L.V.: "SURF: Speeded Up Robust Features", CVIU, vol.110, pp.346-359 (2008).
- 11) Lamdan, Y. and Wolfson, H.J.: "Geometric Hashing: A General And Efficient Model-based Recognition Scheme", Proc. ICCV, pp238-249 (1988).
- 12) 中居知弘, 廣瀬浩一, 岩村雅一: "処理速度とメモリ効率の改善された LLAH によるカメラベース文章画像検索法", 画像の認識・理解シンポジウム論文集, pp. 1252- 1259 (2008).
- 13) Uchiyama, H., Saito, H., Servieres, M. and Moreau, M.: "AR GIS on a physical map based on map image retrieval using LLAH tracking", Proc. MVA, pp382-385 (2009).
- 14) Martedi, S., Uchiyama, H., Enriquez, G., Saito, H., Miyashita, T. and Hara, T.: "Foldable augmented maps", Proc. ISMAR, pp65-72 (2010).