

光学文字認識による音楽演奏インタフェースの提案

卯田駿介^{†1} 馬場哲晃^{†1} 串山久美子^{†1}

光学画像認識技術の発展により、実物体と画像認識を組み合わせた“AR”によるアプリケーションが多く報告されている。ARに利用されるマーカには、オブジェクトの形、色の他、光学マーカと呼ばれる特殊な2次元マーカなどが利用されるが、これらは実物体の物理特徴や印刷パターンであるため、ユーザによって作り出すことはできない。そこで著者らは、ユーザによって手軽にARマーカを作ることができれば、より多くのユーザにとってARシステムにおけるユーザビリティの向上につながると考えた。具体的な手法としては、光学文字認識を利用して、文字をARマーカのように扱うことを検討している。本稿ではそれらの実験段階として、実物体に印刷された文字を利用したシーケンス型電子楽器「Alphabet Sequencer」を制作した。本稿では、「Alphabet Sequencer」のシステムと光学文字認識を利用した音楽演奏インタラクションの展望を報告する。

Proposal of Music Interface using Optical Character Recognition

SHUNSUKE UDA^{†1} TETSUAKI BABA
KUMIKO KUSHIYAMA^{†1}

Recently, many applications by AR(Augmented Reality) have been reported by improvement in image recognition technology. The form of a two-dimensional marker or an object, the color, etc. made a system producer's and cannot make many real objects used for AR by a user. Then, authors think that if AR marker can be easily made by a user, AR system becomes familiar for more people. So, we focus on using OCR(Optical Character Recognition) to make a character as an AR marker. This time, authors try using OCR in the electronic instrument field, and made tangible music sequencer "Alphabet Sequencer" as a prototype system which can sound sequence performance. This paper reports research progress of the present condition of "Alphabet Sequencer."

1. はじめに

画像処理技術の発展により、実物体と光学画像処理を組み合わせたシーケンス演奏可能なARによる電子楽器インタフェースが多く報告されている。

これら電子楽器インタフェースは、従来ディスプレイやマウスなどを利用したGUI(Graphical User Interface)上で操作していた楽器の多重演奏などを、実世界にある実物体で行うことを可能とし、コンピュータ操作が不慣れなユーザにとっても、複雑な作曲・演奏操作を容易にしたインタフェースと言える。

一方、これらARによる電子楽器インタフェースで利用されている光学画像処理のための認識マーカの多くは、色や形、2次元コードといった抽象的なマーカである。そのため、マーカによって表現できる音色の選択肢には限りがあると同時に、マーカの示す音色の役割をユーザが理解することは困難であった。

そこで、本研究では、人々がコミュニケーションするために利用している記号である「文字」に着目し、文字をARマーカとして利用することにした。文字をARマーカに用いることによって、音色の選択肢の制約をなくすと同時に、ユーザにとって、実物体の持つ音色の役割理解向上につながると考えた。

本稿では、研究初期段階のプロトタイプとして制作した、シーケンス演奏可能なテーブルトップ型の音楽演奏インタフェース「Alphabet Sequencer」の概要と考察、光学文字認識を利用した音楽演奏インタラクションについての展望について記述する。

2. 関連研究

(1) 先行研究

著者らはこれまでに、手書き文字や記号を利用して演奏可能なシステムGocen[1]を報告している。このシステムは、ユーザがシステム専用デバイスを用いて、手書き五線譜をなぞり、リアルタイムに演奏できる電子楽器であり、楽器の選択や調の選択には光学文字認識を利用している。Alphabet Sequencerの制作においては、Gocenシステムで得られた光学文字認識に関する技術を利用している。

(2) ARによるテーブルトップ型音楽演奏インタフェース

コンピュータビジョンの普及に伴い、ウェブカメラや赤外線カメラを利用した、ARによるテーブルトップ型音楽演奏インタフェースが多く提案されている。2次元コードをマーカとしたreacTable[2]、色をマーカとしたLEGO STEP SEQUENCER[3]やBubblegum Sequencer[4]、実物体の形状をマーカとしたThe Table is The Score[5]やXenakis[6]など

^{†1} 首都大学東京システムデザイン研究科
Graduate School of System Design, Tokyo Metropolitan University

が代表的な例である。これらの作品は演奏操作が簡単であるが、抽象的なマークを利用しているため、オブジェクトの持つ音の役割をユーザが理解しづらい欠点がある。

3. Alphabet Sequencer について

本システムは 60×60×4mm の予めアルファベットの記述されたカード型オブジェクト、カード型オブジェクトを配置する台座(図1), 200万画素のスタンド型書画カメラ、コンピュータからなる。アルファベットの記述されているカードは、2mm厚の白色アクリル材と半透明乳白色アクリル材を重ねて制作した。アルファベットの記述されている表面は50×50mmの白色アクリル材に黒色で文字列を記述し、半透明乳白色のアクリル材で10mm幅の外枠を形成している。カードを配置する台座は、大きく分けて2つのエリアからなっており、カード16個分を縦4つ、横4つに並べることでできるグリッド状の右側のエリアとカード4つを縦1列に並べられる左側のエリアからなる。右側のエリアはループ演奏の一小節分を表しており、このエリアにアルファベットの記述されたカードを載せることで、どの音高をどのタイミングで鳴らすかが決まる。タイムラインは図2のように、左上から右下に向かって進行する。一方、左側のエリアは楽器の変更や右側エリアに置かれている情報の保持、リセットなどの操作を行うユーティリティーエリアである。

コンピュータソフトウェアは Xcode 上で作成し、openframeworks[7] ライブラリを利用し、ofxMidi, ofxopenCV, などの openframeworks 上で使用できる addon を利用した。音源は Native Instruments 社の Kontakt Player[8] を利用した。



図1 専用の台座
 Figure 1 Original Board

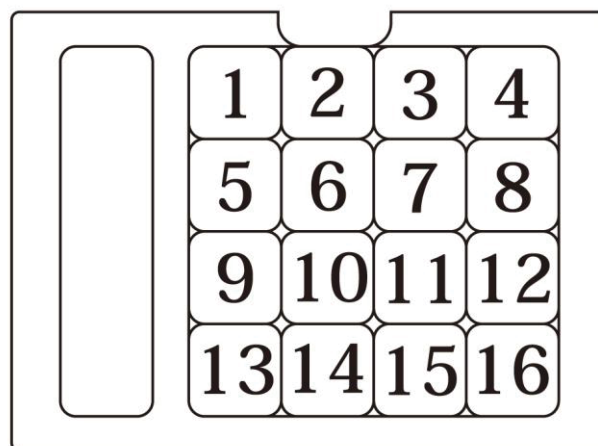


図2 タイムラインの進行順
 Figure 2 Sequence

3.1 ループ演奏

本システムは現状では、1ループ16拍で演奏するシステムとなっている。TENORI-ON[9] や PocoPoco[10] といった多くのシーケンス演奏型の電子楽器では、左側から右側に向かってタイムラインが流れ、右端までタイムラインが進行すると始めの位置にタイムラインが戻るように表現されている。これは現代に一般的な記譜法や文字列が左から右に向かって時間軸が進行していくことからわかるように、人にとって自然な流れだと言える。そこで本システムでも図2のよう、左上から右下に向かってタイムラインが進行していくようデザインした。

3.2 認識における文字の利用

本研究では従来のオブジェクト配置によるループ演奏用音楽インタフェースでは用いられなかった、文字、文字列を AR マークとしている点が特徴である。本システムでの文字の用途を以下に3つ示す。

音高

国ごとによって違いはあるが、古くから音高は文字によって表されてきた。例えば、イタリア音名、ドイツ音名、英語音名、フランス音名、日本音名などがある。本システムでは光学文字認識を利用する上で、凡庸性の高いアルファベットを用いた音名を利用することとした。現行のシステムでは、イタリア音名と英語音名、2種類に対応した(図3)。2種類の音名を利用することで、音楽学習環境の異なるユーザでも、スムーズに利用できることを期待する。

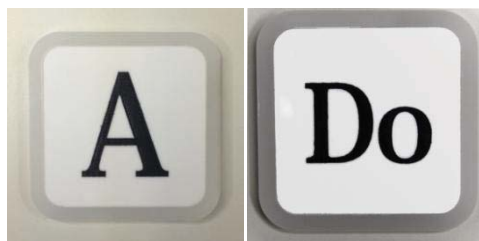


図3 英語音名とイタリア音名の2種類の音階カード

Figure 3 Scale Card

楽器音の選択

楽器の表記には省略記号を用いることとした。省略記号と比べ、楽器本来の名称を用いたほうが、人にとって理解しやすいと考えられるが、楽器本来の名称を用いた場合、楽器ごとに文字数にばらつきがあり、文字の認識率の低下につながると思ったためである。



図4 楽器を示すカード

Figure 4 Instruments Card

ループ中の再生情報の保持, リセット

多くのテーブルトップ型の電子楽器は、従来の古典楽器では出来なかった、「ユーザー人で複数の楽器を同時に再生する」ことが容易となっている。本研究でもその点を踏まえ、楽器ごとに演奏情報を保持していくレイヤーの考え方を採用して、複数の楽器を動的に演奏できるようにした。楽器ごとの演奏情報は、図5の「REC」、図6の「RESET」のカードを用いる。



図5 演奏情報の保持に使用

Figure 5 Record and Reset playing information

4. 考察・課題

本稿で示したプロトタイプ「Alphabet Sequencer」では、

テーブルトップ型の音楽演奏インタフェースで、文字の印字された実物体オブジェクトを操作することによって、楽器や音階の選択などを行うことが可能となった。文字をAR マーカとして利用することで、従来の抽象的なマーカでは限られていた音色の選択肢が、楽器や音階の種類分、表現可能となった。加えて、実物体の示す音色の役割もユーザーにとって理解可能なものとなった。

ただし、プロトタイプを制作し演奏を行ったところ、文字の認識精度には問題があった。文字に入射する光量のばらつきが原因で、文字を正確に認識しないことがある。文字を正確に認識するためには、文字に入射する光量を調整する必要があるため、今後の課題としたい。

5. 展望

手書き文字を利用した音楽演奏インタラクション

文字をAR マーカに利用する利点として、ユーザが自ら手書きでマーカを作りだすことができる点が挙げられる。

今後は、文字をAR マーカに利用した利点を生かし、実物体に文字を印字している現状のプロトタイプの形にとどまらず、手書き文字や複雑な文字列を認識に利用した音楽演奏インタフェース制作を行いたいと考えている。具体的には、手書きのメッセージカードの文字情報を元に演奏を行うオルゴールの制作を予定している。

参考文献

- 1) Tetsuaki Baba, Yuya Kikukawa, Toshiki Yoshiike, Tatsuhiko Suzuki, Rika Shoji, Kumiko Kushiyama, and Makoto Aoki, "Gocen: a handwritten notational interface for musical performance and learning music.", In ACM SIGGRAPH 2012 Emerging Technologies (SIGGRAPH '12). ACM, New York, NY, USA, Article 9, 1 pages. DOI=10.1145/2343456.2343465 <http://doi.acm.org/10.1145/2343456.2343465>
- 2) Sergi Jordà. 2010. The reactable: tangible and tabletop music performance. In CHI '10 Extended Abstracts on Human Factors in Computing Systems (CHI EA '10). ACM, New York, NY, USA, 2989-2994. DOI=10.1145/1753846.1753903 <http://doi.acm.org/10.1145/1753846.1753903>
- 3) Guido Lorenz, "Lego Step Sequencer", ADVANCE HACKATHON, 2010
- 4) Hannes Hesse, Andrew McDiarmid and Rosie Han, "Bubblegum Sequencer", De-Bug Magazine Nr. 135, 2009
- 5) Levin, G. "The Table is The Score: An Augmented-Reality Interface for Real-Time, Tangible, Spectrographic Performance." Proceedings of the International Conference on Computer Music 2006 (ICMC'06). New Orleans, November 6-11, 2006.
- 6) Markus Bischof, Bettina Conradi, Peter Lachenmaier, Kai Linde, Max Meier, Philipp Pötzl, Elisabeth André, "XENAKIS - Combining tangible interaction with probability-based musical composition", Proceedings of the Second International Conference on Tangible and Embedded Interaction (TEI'08), Feb 18-20 2008
- 7) openFrameworks community. openframeworks.com

- 8) Native Instruments. Kontakt player.
- 9) [http://www. native-instruments. com/](http://www.native-instruments.com/)
- 10) ヤマハ株式会社, “TENORI-ON”
- 11) 金井隆晴, 菊川裕也, 鈴木龍彦, 馬場哲晃, 串山久美子,
"PocoPoco: 実物体の動きを利用した楽器演奏インタフェース", 情
報処理学会論文誌, 一般社団法人情報処理学会, Vol. 53, No. 3,
pp1050-1060, 2012