

デモンストレーション：若手による研究紹介IV

浜中 雅俊	科学技術振興機構さきがけ研究員
竹川 佳成	大阪大学大学院情報科学研究科
岩井 憲一	滋賀大学教育学部
高橋 直也	早稲田大学大学院理工学研究科
中野 倫靖	筑波大学大学院図書館情報メディア研究科
大石 康智	名古屋大学大学院情報科学研究科
糸山 克寿	京都大学大学院情報学研究科
北原 鉄朗	京都大学大学院情報学研究科
吉井 和佳	京都大学大学院情報学研究科

あらまし 本デモセッションでは、音楽情報処理の研究分野における若手研究者のさらなる発展に向けて、若手による研究事例をデモンストレーション形式で紹介する。

Demonstrations: Introduction of Research by Young Researchers IV

Masatoshi Hamanaka	PRESTO, Japan Science and Technology Agency
Yoshinari Takegawa	Graduate School of Information Science and Technology, Osaka University
Kenichi Iwai	Faculty of Education, Shiga University
Naoya Takahashi	Waseda University
Tomoyasu Nakano	Graduate School of Library, Information and Media Studies, University of Tsukuba
Yasunori Ohishi	Graduate School of Information Science, Nagoya University
Katsutoshi Itoyama	Graduate School of Informatics, Kyoto University
Tetsuro Kitahara	Graduate School of Informatics, Kyoto University
Kazuyoshi Yoshii	Graduate School of Informatics, Kyoto University

Abstract Toward further progresses of young researchers in the field of music information processing, we introduce case studies of demonstrations.

はじめに

浜中 雅俊

「デモセッション」は今回で4回目の開催となる。毎年1回夏の音楽情報研究会で開催することが恒例となっていたが、本年度は今回の研究会の開催地である関西の若手研究者らの強い希望に応じて、前回の8月に加えて、10月も開催することとなった。

前回8月のデモセッションは、イブニングセッションの形式での初の開催となった。結果的には、研究会の口頭発表件数が非常に多かったため、夕食後の時間を有効に活用するための措置であったが、企画段階でドリンクありでのデモセッションを検討していた我々にとっては願っても無い機会であった。無事終了することができ、

ご協力いただいた皆様感謝する次第である。今後も機会があれば、デモセッションをイブニングセッションの形で開催していきたいと思っている。

前回のデモセッションで行ったもう一つの試みは、研究会に参加できなかった人にもデモの内容が伝わるようデモを写真や動画で記録し配信することであった。現在、若手企画のメーリングリストにご参加いただいている方限定で、公開している¹。多くのデモが同時に行われている中での撮影であったため、手振れや雑音などの問題はあったものの、参加できなかった方々にデモの雰囲気伝わったことには意義があったと思っている。今後は、三脚を用意したり指向性の高いマイクを用意する

¹ 興味のある方は「はじめに」の末尾に案内があるのでご連絡ください。

などの対応を行っていききたい。ただし、写真や動画、そしてこの原稿によって、デモをご覧になれなかった皆様に伝えられる内容は非常に限られており、是非毎回デモセッションをご覧になっていただけることを願っている。

さて、今回のデモセッションは前回のデモセッションからわずか数ヶ月の間を置いての連続開催ということで、デモの出展数が少なくならないか、いささか心配であったが、8件の応募をいただき、前回に引き続き、質・量ともに充実したデモセッションとなるであろう。さらに今回のデモセッションでの新たな試みとして、当日の飛び入り参加枠を用意した。これは、研究会の当日にデモやデモビデオなどを見せられる飛び入り参加者を募り、デモ会場のスペースを使って我々と同じように発表していただくものである。当日に研究会で口頭発表する方や、デモ発表をしてみたいと思っていたがこれまで機会を逃していた方、若手が多いという理由で発表を躊躇していたシニアの方々など、是非この機会にデモ発表を検討していただければと思う。

我々がデモセッションを開始した当初の目的の一つは、未完成の研究をみんなで育てて行く場をつくるということであった。今回のデモの中には前回のデモセッションと2回連続してのデモとなったものもある。前回のデモで行われた議論がどのように反映されているのか非常に興味深いところである。さらに次の展開について、みんなで議論していけることを楽しみにしている。また、この「デモセッション」自体もまだまだ未完成である。デモセッションを今後どのように発展させていくかについても皆様から是非ご意見をいただきたいと思っている。

デモセッションでは興味深いシステムをその場で見たり触ったりでき、それ自体非常に楽しいものである。しかし、それと同じように重要なのは、人と人との交流ができる点である。通常の口頭発表では、発表はしたものの、それっきりということがあってはならないかと思う。これに対し、デモセッションでは出展者がデモ展示物の説明をするだけでなく、往々にして見学者の側からアドバイスを受けることがある。また、その過程でお互いに名刺交換をしたり研究内容を紹介することもある。それが共同研究を開始するきっかけとなることもありえるだろう。そのような交流が、本研究会の発展の一助となることを願っている。

Sigmus 若手企画メーリングリスト 参加者募集

Sigmus 若手企画メーリングリストは、これまでのデモセッションに参加いただいた方々に登録いただいているもので、デモセッションの事務連絡および、今後の若手企画の相談、連絡等を行っています。年齢等制限ありません。今後の企画に参加したいと思っている方は是非メーリングリストに加わってみませんか。

問い合わせ先：

浜中雅俊 (m.hamanaka[at]aist.go.jp [at] @)

Bluetooth-Midi Adapter の試作

浜中 雅俊, 上田 健太郎

MIDI 信号の双方向通信を可能とする Bluetooth-MIDI Adapter を試作した。従来、ステージ上で離れた位置にある MIDI 機器を相互に接続するためには、長い MIDI ケーブルで両者を接続しなければならないという煩わしさがあった。

Bluetooth-Midi Adapter は、2 台の MIDI シーケンサなどにそれぞれ親指大のインタフェイス(図1)を取り付けるだけで、相手の演奏を Bluetooth を介して自分の MIDI シーケンサで聴くことを可能とする。つまり、Bluetooth-Midi Adapter 一对で一本の MIDI ケーブルの役目を果たす。現在、Bluetooth 搭載の PDA と MIDI 音源を接続するためのアダプターも製作中である。



図 1: 試作した Bluetooth-Midi Adapter

このような Bluetooth を用いた MIDI インタフェイスは、2001 年の NAMM ショー(楽器展示ショー)で Midiman 社(現 M-Audio 社)により参考出品されていたが、市販には至らず、また、その詳細は明らかにされていなかった [1]。

今回の展示では、GridConect 社の Bluetooth RS232 Serial Pair Firefly Pair を用いて、MIDI シーケンサ(YAMAHA QY100)の TO HOST ポートから入出力される 38400bps の MIDI 信号を Bluetooth に変換した。

レイテンシの測定は行っていないが、リバーブを含む音色では遅延がほとんど感じられなかったが、リバーブを切るとわずかに遅延が感じられた。Bluetooth に変換する際のバッファを小さくするなどの改良について検討していきたい。

今後はビットレートの変換を行うことで、31250bps の DIN5pin の MIDI ポートにも接続できるようにしていき、ギタリストがエフェクターの音色を切り替えるときに用いる MIDI フットコントローラをワイヤレス化するなど実現していきたい。

参考文献

[1] <http://namm.harmony-central.com/SNAMM01/Content/Midiman/PR/Bluetooth-MIDI.html>

TEMPEST: 音楽的表現が可能な文字入力支援システム

竹川 佳成, 寺田 努, 西尾 章治郎

近年, 計算機の普及に伴い, 電子メール, ブログ, チャットなどテキストベースのコミュニケーションが活発に行われている。また, 感情や込めたい思いをより良く表現し他者に伝えるために, 写真, ムービー, 効果音, 音楽などさまざまなコミュニケーションメディアが提案されてきた。一方, ピアニストは鍵盤楽器を使って感情や思いを巧みに表現している。したがって, 文字入力に演奏者の音楽的な表現力を付加できれば, 感情や思いをより豊かに表現・伝達できると考えられる。そこで本研究では, 鍵盤を使って文字入力に音楽表現を付加できる文字入力システム TEMPEST (TExt input and Musical PErforming SysTem) を提案する。TEMPEST の文字入力方式は, 楽器演奏による感情伝達を考慮した設計になっており, ユーザはあたかも演奏しているかのように文字入力できる。

システムの設計

TEMPEST におけるキーボード上の鍵は, 打鍵時に文字と楽音を生成する文字入力鍵と文字入力に用いない演奏鍵の 2 種類がある。提案する文字指定方式は, 各文字に対して使用する鍵が決まっているため, 文字に依存した芸術性に乏しい音楽が出来上がってしまう。この問題を解決するために, 文字入力鍵に割り当てられているノートナンバーを演奏領域で弾いているノートナンバーに応じて変化させる。具体的には, 演奏領域で弾いている音と不協和音にならない音集合を導出し, 文字入力領域にそれらの音を敷き詰めることで違和感のない文字入力演奏を実現する。また, TEMPEST は打鍵の強さや音色など演奏表現に合わせて文字の大きさや色などを修飾できる付加機能を持ち, ユーザはよりエンタテインメント性高い文字入力演奏を楽しめる。

図 1 に TEMPEST の利用図を示す。今回のデモセッションでは, TEMPEST を実際に体験していただくことで, 演奏で文字を入力することの可能性について議論したいと考えている。

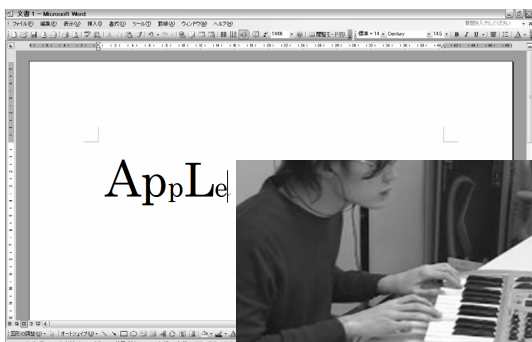


図 1: TEMPEST の利用図

画像の色彩情報に基づく対話型作曲システム MUSCAT の紹介

岩井 憲一, 外村 麻純, 藤井 由香

感性情報を利用したアルゴリズム作曲手法は, それを利用しないアルゴリズム手法とは異なり, 聴き手の感性に近い楽曲が作成できると考えられている。筆者らは, 感性を取り扱う手法として画像の色彩情報を採用し, この情報に基づいてユーザと対話的に作曲を行っていくシステムである MUSCAT (MUSic Composer bAsed on color information of picTures) [1] を構築した。

(1) MUSCAT の楽曲作成手法

具体的な楽曲作成手法を以下に示す。楽曲構造には, 例えばポップスでは「A-A-B-B」や「A-A-B-A」のようなものがあることから, その「A」や「B」に該当する画像として, まずユーザが JPEG 画像を 2 枚入力する。各々の画像に対してシステムが入力画像の画素の中から特徴色を HSB 値の形で抽出する。現時点では最頻のものを特徴色としている。次に, ファジィ推論によって, その特徴色が属する色カテゴリーの判定と, 関連する感性形容詞の同定を行っていく。

感性形容詞が同定できれば, それに応じて楽曲作成に必要な楽曲素材として, 楽曲構造・楽曲中のコード進行・リズムパターン・音色等を各々のデータベースから検索し, それらを元に楽曲の作成を行う。できあがった楽曲候補群はシステムによって候補の選定が行われるが, 現時点では必要な知識の獲得と楽曲生成時における組合せ爆発回避の観点から, 楽曲素材導出の際に随時ユーザに対して図 1 のような評価入力ウインドウを用いて確認を求める形式で対話的に運用されている。なお, 図中の MPT とは本研究において Melody Pattern Template と名付けられたメロディ用リズムパターンである [1]。



図 1: 評価入力ウインドウ

(2) おわりに

今後は, MUSCAT で作成される楽曲の質を高めるために, 素材となるフレーズの充実や, 感性形容詞と色彩情報との関連性をさらに進めていく予定である。

参考文献

- [1] 岩井 憲一, 中川 早織: 画像の色彩情報に基づく対話型作曲システム MUSCAT について, 情報処理学会第 68 回全国大会講演論文集, Vol.2, 2H-2, pp.115-116, 2006.

共鳴部駆動による箏のエレキ化

高橋 直也, 橋本 周司

(1) はじめに

近年, 邦楽の海外進出やセッションの多様化により, 箏においてもドラムやエレキベースなどの大音量楽器とセッションすることが増えている. その際, 弱音量楽器の箏の音量を増幅し, バランスをとる必要がある. 多くのライブでの箏のマイキング位置は, 箏の音量が最も大きいサウンドホールの真下に置かれることが多い. しかしながら, 箏はサイズが大きく, 弦や共鳴部全体で響く楽器であるため, 本来前方で聴く箏の音色とは異なってくる. またこのようなマイキングでも, 他の大音量楽器の音がマイクロフォンに侵入し, 箏の音量のみを増幅できず, 箏の音が聴こえにくかったり, バランスを取れなかったりする. また, ハウリングが起こり十分な音量を出せないことも少なくない.

このように弱音量楽器である箏は音質劣化, 外部音侵入, ハウリングの問題から十分な音量増幅と音色保持を両立することが難しい. ここではこれらを同時に解決する手法として, 箏の共鳴部を駆動することにより音を発生させ, 音量増幅を図る手法を提案する.

(2) 提案手法

図1に提案手法の概念図を示す. 弦はオシレータ, 共鳴部は音色を生成するフィルタに対応する. 矢印は電気, 音響等の信号を表す. 弦の振動を増幅して共鳴部に伝えることにより音量増幅を図る. 音を生成するフィルタに共鳴部そのものを用いているため, 音質を変化させることなく音量増幅ができると考えられる. また空気振動を利用しないピックアップを用いることで外部音侵入を低減できると考えられる.



図1: 提案手法の概念図

(3) エレキ箏の作製

提案手法の実現方法として, 図2に示すようなエレキ箏を作成した. ピエゾピックアップにより弦の振動を拾い, ハウリング防止及びピエゾの特性補正用のイコライザを通してアンプに入力し, 共鳴部内部に取り付けたスピーカから音を再生することにより共鳴部を駆動させ音量増幅を図る.

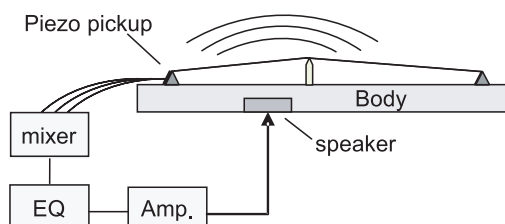


図2: エレキ箏の構成図

歌唱力向上支援システムの試作:

楽譜を用いないリアルタイム音程評価提示機能

中野 倫靖, 後藤 真孝, 平賀 謙

歌唱力向上支援システムの一環として, 人が歌唱力評価に用いる特徴 [1] のうち, 音程の正しさをリアルタイムに点数化して提示するシステムのデモを行う.

(1) 使い方

ユーザは, リアルタイムに提示される音程評価結果を見ながら, 自分の好きな歌・好きな箇所を歌い続け, 評価が低いと判断した瞬間にその付近を歌い直したり, 急に違う曲を歌ったりする使い方ができる.

(2) 実現方法

前回の試作システム [2] は, F_0 軌跡から算出した *semitone stability* (以下, $P_g(F, t)$ と呼ぶ) [3] を音程評価の指標として提示したのみで, 分かりやすさの点で問題があった. そこで本システムでは, $P_g(F, t)$ の短時間平均のピークの鋭さを点数化して順次提示する (図1).

F_0 が, ある基準ピッチのもとで 100cent (半音) の整数倍で遷移していれば, $P_g(F, t)$ の短時間平均は鋭い1つのピークを持ち, その鋭さを評価することで音程の正確さを評価できる. 現在の実装では, 2秒の矩形窓を 200msec ずつシフトさせて $P_g(F, t)$ を算出し, その 10フレーム (4秒) の短時間平均を利用した.

点数はその短時間平均の2次モーメントを利用して求め, F_0 が 50cent 単位で遷移する場合を 0点, 100cent 単位で遷移する場合を 100点として設定した. また「うまい」歌唱者と「へた」な歌唱者それぞれ 300サンプルずつから算出した点数の平均値, 標準偏差, 最大値, 最小値を同時に表示した.

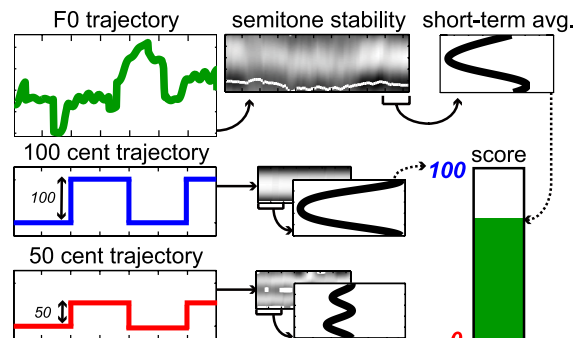


図1: $P_g(F, t)$ の短時間平均からの点数化

謝辞 本デモは, RWC 研究用音楽データベース (ポピュラー音楽), AIST ハミングデータベースを利用しました.

参考文献

- [1] 中野倫靖, 後藤真孝, 平賀謙: 原曲を知らない聴取者による無伴奏歌唱の歌唱力評価, 音楽知覚認知学会 平成 18 年度春季研究発表会, pp.11-16, 2006.
- [2] 浜中雅俊 他: デモンストレーション: 若手による研究紹介 III, 情処研報 2006-MUS-66-10, Vol.2006, No.90, 2006.
- [3] T. Nakano, M. Goto, Y. Hiraga: An Automatic Singing Skill Evaluation Method for Unknown Melodies Using Pitch Interval Accuracy and Vibrato Features, in *Proc. Interspeech2006*, pp.1706-1709, 2006.

音声だけでシームレスにハミング検索と曲名検索が可能な楽曲検索システムの提案

大石 康智, 後藤 真孝, 伊藤 克亘, 武田 一哉
メロディの歌唱(歌声)と曲名の読み上げ(朗読音声)で検索可能な楽曲検索システムを提案する [1]. 本システムは, 歌声と朗読音声を自動的に識別する技術が組み込まれているため, ユーザは入力音声の発声スタイルを切り替えるだけで, シームレスに楽曲を検索することが可能である. 図 1 のように音声識別器, ハミング検索, 音声認識の 3 つの主なモジュールから構成され, 各モジュール間は, 音楽用通信プロトコル RMCP (Remote Media Control Protocol) [2] にしたがって, パケットを送受信することにより, 分散して処理が進められる. 以下に各モジュールの役割を示す.

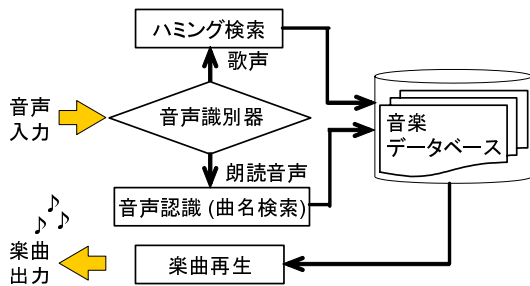


図 1: 楽曲検索システムの構成

(1) 音声識別器

歌声と朗読音声の自動識別を実現するために, 音声信号の短時間スペクトル, また基本周波数の時間変化を識別尺度として利用する. その結果, 2 秒の音声信号に対して 87.3% の識別性能を実現した [3].

(2) ハミング検索

歌声と識別された入力音声から基本周波数を抽出し, メロディの変化を記述するパターンに変換する. 音楽データベース中の各楽曲のメロディに対しても同様のパターンに変換し, パターン間の連続 DP に基づいて, 類似度の最も高い楽曲を再生する.

(3) 音声認識(曲名検索)

朗読音声と識別された入力音声に対して音声認識を行い, 言語情報を抽出する. 「[アーティスト]の[曲名]を聴かせてください」や「[曲名]が聴きたい」というように複数の認識用文法を用意し, 認識された[曲名]に基づいて楽曲を再生する.

参考文献

- [1] 大石康智, 後藤真孝, 伊藤克亘, 武田一哉: 音声だけでシームレスにハミング検索と曲名検索が可能な楽曲検索システム, 情報処理学会論文誌, 2006-MUS-67-2, 2006.
- [2] 後藤真孝, 根山亮, 村岡洋一: RMCP: 遠隔音楽制御用プロトコルを中心とした音楽情報処理, 情報処理学会論文誌, Vol.40, No.3, pp.1335-1345, 1999.
- [3] 大石康智, 後藤真孝, 伊藤克亘, 武田一哉: スペクトル包絡と基本周波数の時間変化を利用した歌声と朗読音声の識別, 情報処理学会論文誌, Vol.47, No.6, pp.1822-1830, 2006.

混合音中のパートの音量を操作可能なオーディオプレーヤー

糸山 克寿, 後藤 真孝, 奥乃 博

音楽を聴くためのツールやインターフェースには様々なものがあり, 特にグラフィックイコライザやサウンドカードのエフェクト機能はユーザが新鮮な気分で音楽鑑賞を楽しむことを可能としてきた. しかし, これらのツールは必ずしもユーザの要求を満たすものではない. ユーザが欲しい機能は周波数帯域に対する操作ではなく「ベースを抑制して, ギターを強調したい」「ボーカルだけにエフェクトをかけたい」といった, 楽器パートごとに対することが多いのではないだろうか. 本研究で製作したオーディオプレーヤーはこのようなユーザの要求を満たすべく, 楽曲に含まれるそれぞれの楽器音が構成するパートに対して, 音量などを自由に操作することを可能にした. 設計したインターフェースを図 1 に示す.

(1) 音源分離とモデル適応

楽器パートごとに対する操作を行うためには, 楽器パートごとの音響信号を分離する必要がある. 分離のため, 本システムでは楽譜 (MIDI ファイル) を用いる. 分離の手順は以下ようになる.

1. MIDI ファイルから抽出した単音を MIDI 音源で録音して, テンプレート音を作成する.
 2. テンプレート音から楽器音モデルを作成し, このモデルを使って各パートの音響信号を分離する.
 3. テンプレート音と楽曲の音色の違いを吸収するため, 楽器音モデルを分離音に適応させる.
 4. 適応させた楽器音モデルを用いて, 再び分離を行う.
- あらゆる楽器音を扱うためには, 調波構成成分と非調波成分の一方だけでは不十分で, 双方を考慮する必要がある. そのため, 楽器音モデルには調波構成成分 (パラメトリックに定義) と非調波成分 (ノンパラメトリックに定義) を併用する新たなモデルを用いている.

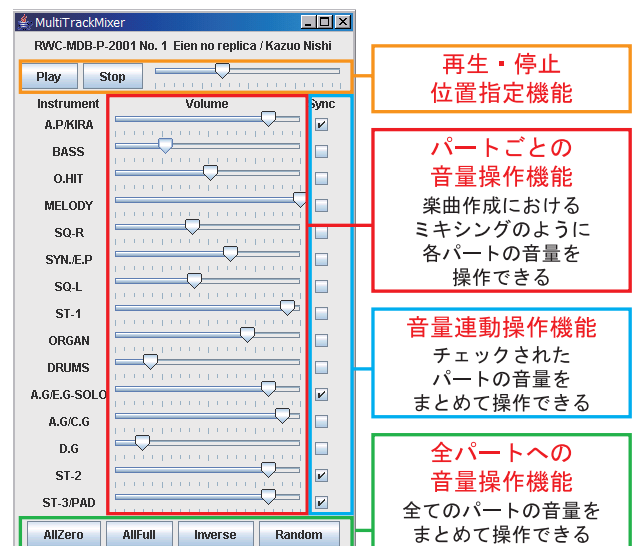


図 1: プレーヤーのインターフェース

Instrogram による類似楽曲検索

北原 鉄朗, 後藤 真孝, 奥乃 博

Instrogram を用いて楽器構成に基づいた類似楽曲検索を行うシステム [1,2] を紹介する.

(1) Instrogram とは

Instrogram は, 図 1 に示すような, それぞれの楽器が音を鳴らしている確率(楽器存在確率)を時間・周波数平面上に表したものである. 従来の楽器音認識と異なり単音(一音符の音)を処理単位としないので, 単音の抽出が不要という特長がある.

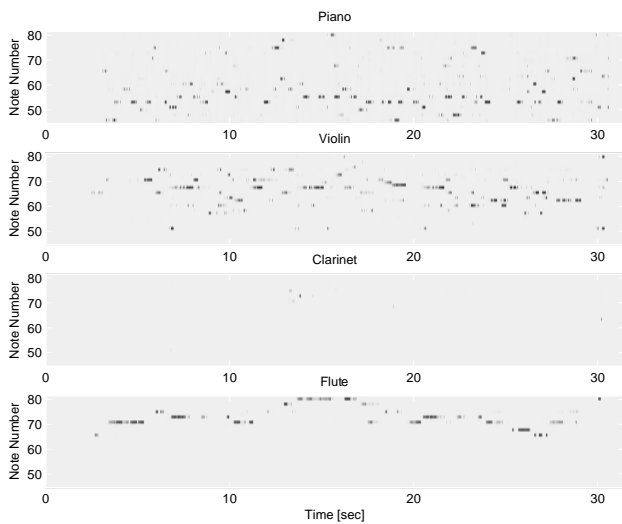


図 1: Instrogram の例(Flute, Violin, Piano による三重奏)

(2) 検索システム概要

Instrogram 同士の類似度を DP マッチングで求めることにより, Instrogram(つまり楽器構成)の類似度に基づいた音楽検索を実現した(図 2). プロトタイプが <http://winnie.kuis.kyoto-u.ac.jp/~kitahara/instrorgam/> から入手できるので, 一度試していただきたい.

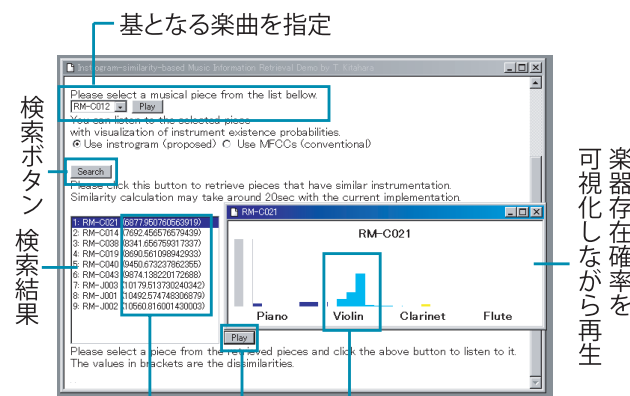


図 2: Instrogram 間の類似度に基づく音楽検索

参考文献

- [1] 北原 他: “Instrogram: 発音時刻検出と F0 推定の不要な楽器音認識手法”, 情処研報, 2006-MUS-66, pp.69-75 (2006).
- [2] T. Kitahara et al.: “Musical Instrument Recognizer “Instrogram” and Its Application to Music Retrieval based on Instrumentation Similarity”, Proc. ISM 2006, in press (2006).

おわりに

吉井 和佳

今回で早くも 4 回目の開催となった若手デモセッションには 8 件ものシステムの応募があり, 力作揃いとなった. 前回のデモセッション開催(10 件発表@夏のシンポジウム)から 3ヶ月弱しか経過していないことを鑑みると, 若手研究者の発表意欲が高いことの現れであろう. 前回の発表者でもシステムを洗練させて, あるいは全く新しいシステムを開発したので今回も発表したいという意欲が高かった. 日程の都合上, 発表を断念せざるを得なかった方々も多くいたことも記しておく. 次回のデモセッションの参加は大歓迎であるので是非参加をお願いする次第である. このように音楽を通じた研究の輪は着実に広がっていると感じる.

本企画は自分のシステムを実演したいという要求を満たす場を提供してきただけでなく, 発表のやり方に関しても知恵を出し合って工夫を重ねてきた. 例えば前回は研究発表会プログラムが 2 日間に渡り, 宿泊を前提をしていたので, デモセッションはイブニングセッション(要は宴会セッションである)と銘打って夕食後に開催した. 今回は研究発表会前日の CrestMuse シンポジウム² と共催であり, 2 日間合わせると多数の最新音楽システムを見て回ることができる. このように本企画は回を重ねるごとに回ごとの特徴がはっきりと打ち出されるようになり, 発表者・見学者ともに毎回違った楽しみ方ができるようになった. また, 未完成のシステムでも気軽に発表することを奨励してきたことは特筆に価する. 研究者は論文には書かない(書けない)けれどシステム開発のための有用なノウハウを持っていることが多いため, 発表者と参加者が互いの生の苦労を分かち合いながら議論を行うことは有意義である. 現在の他分野の状況を見ると研究会レベルでデモセッションを開催することは稀であるが, 先陣を切って興味を持つ人の裾野を広げていこうとする努力はまさにフロンティア領域と名乗るに相応しい.

ただし, システムを実際に使ってもらわないとおもしろさが伝わらない, と諦めてしまっては試合終了である. 使わせなくてもおもしろさを伝える・広めるための努力を惜しんではならない. そのための一つの手段はシステム紹介のための WEB ページを開設することである. アプリケーションとして配布できれば理想的であるが, 研究段階のプロトタイプ実装であっても, 動作の様子が収められたビデオを公開しておくことは重要であると思われる. そうしておけば, 論文を読んだ読者(査読者かもしれない)は内容を直ちに正しく理解してくれるであろう. 現代人には時間がないので「百聞は一見にしかず」なのである. もちろん論文は「百聞」しなければならないような冗長性がないように執筆する必要がある. 今後, デモセッション自体を盛り上げるだけでなく, このようなことにも注意を払う必要があるだろう.

² <http://www.crestmuse.jp/>