

デモンストレーション: 音楽情報処理の研究紹介 XII

深山 覚¹ 石倉 和将² 和泉 諒³ 岩井 憲一⁴ 植村 あい子² 奥村 健太⁵ 加庭 輝明⁶ 河原 英紀⁷
小池 宏幸⁸ 須之内 元洋⁹ 竹川 佳成¹⁰ 土屋 政人¹² 中村 友彦¹² 中村 亮介¹³ 濱崎 雅弘¹
ベンジャミン エバンズ⁹ 堀内 俊治¹⁴ 松原 正樹¹⁵ 三浦 寛也¹⁶ 栗 孝之¹⁷

概要: 本セッションでは、デモンストレーション形式で、音楽情報処理の研究を紹介する。

解説

深山 覚 (モデレータ)

このセッションはこれから音楽情報科学の分野に飛び込む学生の方などへの水先案内の意義があると聞いた。したがって本稿では今回の発表者のデモ内容概略に触れつつ音楽情報科学の分野でどのような研究があるのか解説を試みたい。

はじめに音楽創作に関わる研究である。1950年代の大型計算機が登場して間もない頃からの研究領域である。計算機は新しい作曲スタイルを開拓し、多くの新しいコンセプトと音楽が生まれている。岩井氏は感性情報の作曲プロセスへの導入を行っている。栗氏と松田氏は音楽処理言語Maxにリアルタイム映像処理を加えるオブジェクト群を作曲家達に提供している。音楽の専門家でない人のためには、ベンジャミン氏が作曲実技の知見を計算機に実装する研究を行っている。

楽器研究の歴史も古い。現在では、従来からの楽音合成の研究に加え、楽器のインタラクティブ性に着目した研究が行われている。小池氏の紹介する電子楽譜プラットフォームには楽器練習ツールが付属している。竹川氏はPC用のキーボードカバーを利用した新しい演奏用インタフェースを研究している。

人間だけでなく機械も演奏ができないか、と考える人たちがいる。奥村氏は統計モデルを生かして人間の演奏を分析し、その知見を自動演奏に生かしている。中村友彦氏は自分の演奏に合わせて伴奏をする計算機システムを作っている。中村亮介氏は機械を含む複数奏者による協調的なアンサンブル演奏の実現法を研究している。

機械が人間にとって代わるかという点で、長年挑戦されている課題が音からの譜面書き起こしである。これを実現するのが難しい理由の一つは、重なりあう音の分析が難しいからだということがわかってきた。和泉氏はその問題に対して現実的な使用場面を考えた処理速度を重視する解法を提案している。音一つ一つまで行かずとも、和音の解析でも十分挑戦的な課題であり、植村氏が取り組んでいる。もう一つの困難は楽譜と実際に演奏される音のずれである。人間の演奏にはテンポや音量の変動がある。土屋氏はあらかじめ楽譜が分かっているならば、音響信号中の音が楽譜のどの音に対応しているかを解析できるシステムを研究している。

音楽情報科学の問題意識は時代とともに大きく変わって来ている。音楽配信サービスの普及に伴い音楽検索・推薦・楽曲データ可視化の研究が多いに活発化した。石倉氏、濱崎氏の研究はこれに関連している。さらに音楽データをより能動的に加工して楽しむという動きもある。自分の好みにあわせて楽曲に手を加えて聞く技術には中村友彦氏が取り組んでいる。スマートフォン・タブレットの普及も大きな影響がある。小池氏はタブレット型デバイスを用いた電子楽譜閲覧・配信用プラットフォームを提供している。堀内氏はスマートフォン向けのミュージックビデオアプリケーションを開発している。スマートフォンの高機能化はますます身の回りの音の録音・デジタル化を促進させると思われる。須之内氏が取り組んでいる環境音検索は、その中でますます重要になっていくだろう。

また近年は歌声研究が大変盛んである。今回の記念シンポジウムでもスペシャルセッションが開催されている。河原氏などによって音声分野で培われた数理的定式化や信号処理のセンスで実用的な手法が多く研究されている。その他、可聴化は医療現場や製品品質管理から芸術にも通じる魅力的な領域である。直感的なメディアインタラクションが増える今後ますます重要になるであろう分野である。加庭氏、松原氏はこの分野の研究に取り組んでいる。

最後に、音楽分野で培われた技術が、音楽以外の場面で有効利用される快挙を狙う動きもある。三浦氏の研究は、そもそも音楽理論であるGTTM理論を議論の分析につかっってしまうという、音楽の枠から飛び出る意欲的な研究である。

音楽情報科学が異分野から見て魅力的な分野に映るよう、私自身も分野を盛り上げるべく研究を頑張っていきたい。

¹ 産業技術総合研究所
² 早稲田大学大学院基幹理工学研究科
³ 東邦大学大学院理学研究科
⁴ 滋賀大学教育学部
⁵ 名古屋工業大学
⁶ 筑波大学大学院システム情報工学研究科
⁷ 和歌山大学システム工学部
⁸ プラスアド株式会社
⁹ 北海道大学大学院
¹⁰ 公立はこだて未来大学システム情報科学部
¹¹ 東京大学大学院情報理工学系研究科
¹² 東京大学大学院情報理工学系研究科
¹³ 横浜国立大学大学院環境情報学府
¹⁴ KDDI 研究所ユーザインタフェースグループ
¹⁵ 筑波大学図書館情報メディア系
¹⁶ 公立はこだて未来大学大学院
¹⁷ ランカスター大学ランカスター現代アート研究所

自動で曲名を教えてくれる楽曲同定機能付きメディアプレイヤー

石倉 和将, 植村 あい子, 甲藤 二郎

メディアプレイヤーで再生されたカバーソングなどの類似楽曲をそれらの対応する原曲と同定するシステムを紹介する。近年スマートフォンアプリとして楽曲同定システムは多くリリースされているが、それらではカバーソングやライブバージョンの同定は難しい。そのため我々はそれらのアプリが用いる手法に比べ抽象度が高く、カバーソング楽曲同定によく用いられるビート-クロマ特徴量を利用した楽曲同定システムを構築した。

1. アプリケーション概要

図 1 に今回紹介する楽曲同定システムの GUI を示す。本アプリケーションでは、組み込まれたメディアプレイヤーで再生されている動画・音楽の音声信号から特徴量を抽出・解析することで楽曲の曲名を表示する。また、楽曲同定処理には筆者らが過去に提案した手法 [1] に改良を加えたもののほか、カバーソング楽曲同定に関わる複数の手法を利用することができる。

2. 楽曲同定のアルゴリズム

本システムにおける楽曲同定は次の流れで行われる。

- (1) 音響信号を数十秒毎に切り出す (クエリ生成)
 - (2) クエリからビート-クロマ特徴量を抽出する
 - (3) データセット内の原曲のビート-クロマ特徴量とマッチング (相互相関) を行う
 - (4) マッチング結果を基に同定楽曲を表示する
 - (5) 原曲の楽曲長を基に次クエリに重み付け処理を施す
- 1~5 の処理をメディアプレイヤーの再生が終わるまで繰り返し行う。クエリ長を数十秒としているため連続するクエリは何度か同じ同定結果を出力することが望ましく、また、ユーザーがライブ DVD を見ている場面というものをひとつの用途として想定しているため、同じ楽曲が何度も演奏されることは稀であるというライブ特有の特徴への対応のために類似度スコアへの重み付け処理を行っている。

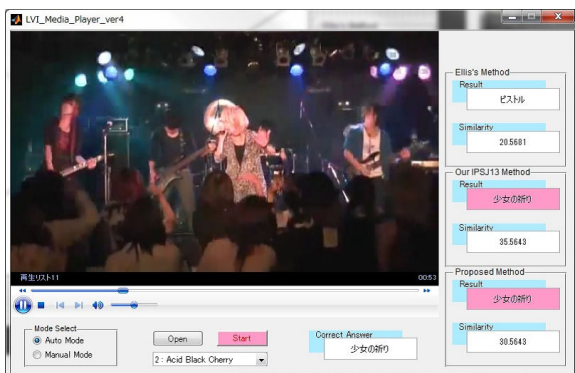


図 1 アプリケーション GUI

参考文献

- [1] 石倉和将ほか, : ライブにおける楽曲の遷移を考慮したリアルタイム楽曲同定システム, 情報処理学会研究報告, 2013-MUS92-1 pp.1-6 (2013).

多重音に対する音高・音源数の高速推定法

和泉 諒, 松島 俊明

多重音に対して、音高と音源数を高速に推定するシステムの紹介を行う。本研究の方法は、自動譜めくりシステムに使用することを目的としている。そのため、推定精度よりも処理速度を重視した方法を採用している [1]。

1. 推定アルゴリズム

本方式による多重音の音高推定法は、ハーモニッククラスタリングで用いられている方法を基本としているが、あらかじめ複素スペクトル内挿法で基本周波数候補を絞ることで、ハーモニッククラスタリングでの反復計算回数を抑え、高速推定を実現している。また音源数の推定については、基本周波数候補に倍音に対応するピッチが追加された場合は評価関数の値の変化が小さくなる性質があることを利用している。以下、本方式の基本的なアルゴリズムを示す。

- (1) 短時間フーリエ変換により得られたパワースペクトルから一定以上のパワーを持つ周波数ピーク位置を複素スペクトル内挿法により抽出する。
- (2) 推定音源数 $K = 1$ に設定する。
- (3) 一番低いピークから上位 L 個 ($L > K$) の周波数ピークを基本周波数候補として選出する。
- (4) L 個のピークから K 個を選択して得られる ${}^L C_K$ 組の基本周波数の組み合わせの集合を作成する。
- (5) 上で得られた全ての組み合わせに対してハーモニッククラスタリングによる評価関数の値を計算し、その最小値 $D(K)$ とその時の基本周波数候補 μ^K を求める。
- (6) $K = 1$ ならば $K \leftarrow K + 1$ として 3. に戻る。
- (7) $2 \leq K$ のとき、 $D(K)/D(K-1)$ が 1 に近い値であれば、音源数を $K-1$ 、基本周波数を μ^{K-1} に決定する。そうでなければ $K \leftarrow K + 1$ として 3. に戻る。

2. 処理例

単音から 6 和音までが混在した音源に対する音高・音源数の同時推定結果を図 1 に示す。ノート PC (Intel CORE2Duo 2.66GHz CPU, 2.96GB RAM) で処理した場合、4 和音程度であれば処理落ちせずに高速推定することが出来た。

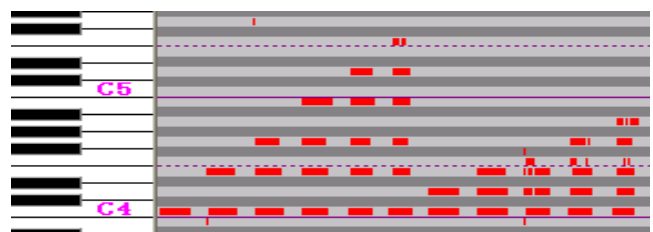


図 2 単音・6 和音の推定例

参考文献

- [1] 安部翔, 小田弘良, 松島俊明:自動譜めくりのための多重音の音高・音源数の推定法の高速化, 情報処理学会研究報告, 2011-MUS92-1 pp.1-6 (2011)

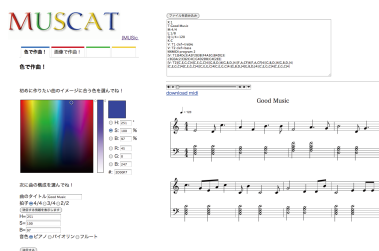
色彩情報に基づく自動作曲 Web アプリケーション MUSCAT の紹介

岩井 憲一

筆者はこれまでに、感性情報を利用した自動作曲アルゴリズムの実現を目指して、画像の色彩情報に基づく対話型作曲システム MUSCAT (MUSIC Composer bAsed on color information of picTures) [1][2] を構築した。今回、作曲プロセスの自動化を含めた機能の洗練化を推し進めるとともに、システムの有用性と可動性の向上を目指して Web アプリケーション化にも取り組み、新たに色彩情報に基づく自動作曲 Web アプリケーション MUSCAT (MUSIC Composer bAsed on color informaTion) を実現した。

1. MUSCAT の楽曲作成手法

図 3(a) に Web アプリケーション化された MUSCAT のインターフェースを示す。まずインターフェース上のカラーピッカーから色を一選択し、この色の HSB 値を抽出して MUSCAT に情報を送信する。MUSCAT はこの HSB 値を元にファジィ推論によってその色が属する色彩カテゴリーの判定とあらかじめ色彩カテゴリーと関連づけられた感性形容詞の同定を行っていく。感性形容詞が同定されれば、それに応じた楽曲編成に必要なコード進行と感性に応じたメロディやリズムのテンプレートをそれぞれのデータベースから抽出して楽曲の合成を行い、図 3(b) のように作成した楽曲をインターフェース上で大譜表の形式で出力する。また、併せて MIDI データにも変換しているため楽曲の再生も可能である。



(a) インターフェース (b) 出力例

図 3 MUSCAT のインターフェースと出力例

2. 利用環境について

MUSCAT は QuickTime をインストールした PC や Mac における Internet Explorer 以外の Web ブラウザ上で利用可能である。また、Tablet PC やスマートフォン上でも大譜表の表示は可能であるが、楽曲再生については、これまでに Android OS 系 Tablet PC における Firefox での動作が確認できている。

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.
- [2] iwai2 岩井 憲一, 外村 麻純, 藤井 由香: 画像の色彩情報に基づく対話型作曲システム MUSCAT の紹介, 情報処理学会第 67 回音楽情報科学研究会デモセッション講演論文集, pp.3, 2006.

異なるビットレートの楽曲を用いたクロマベクトルの解析と和音認識性能評価

植村 あい子, 石倉 和将, 甲藤 二郎

本研究は、音楽情報処理で使用される特徴量クロマベクトルとそれを用いた和音認識について、楽曲の圧縮の影響を調査する。周波数成分を変形させる処理は音楽要素の認識やコンテンツベースの情報検索の結果に影響を及ぼすことが懸念される。本稿では、ビットレートを变化させた入力楽曲の和音認識を行うとともに、客観品質評価を行うアプリケーションを作成した。

1. システムの全体像

我々は、圧縮音源に対して和音認識システムを作成した。図 4 にシステムの GUI 画面を示す。システムの作成には MATLAB の GUI アプリケーション構築ツール GUIDE を使用した。

2. 和音認識システム

本アプリケーションにおいて、楽曲からの自動和音認識は、次の流れで行われる。

- (1) 入力信号からビートごとに特徴量クロマベクトルを算出する。
- (2) 学習データを用いて各フレーム (ビート) に対し HMM/SVM により和音ラベル付けを行う [1]。
- (3) 推定されたラベルと正解ラベルを比較し、和音認識率を算出する。

なお、本アプリケーションでは学習時間削減のため、和音認識用データセットとして提供されている The Beatles, QUEEN, C.King と RWC 研究用音楽データベースよりポピュラー音楽計 307 曲をあらかじめ学習したモデルを保有する。認識する和音は major と minor の 24 和音を扱うこととし、提供される正解ラベルは人手で 24 和音に分類した。音質の評価には ITU-R BS.1387-1 で定義されている PEAQ(perceived evaluation of audio quality) を用いる。

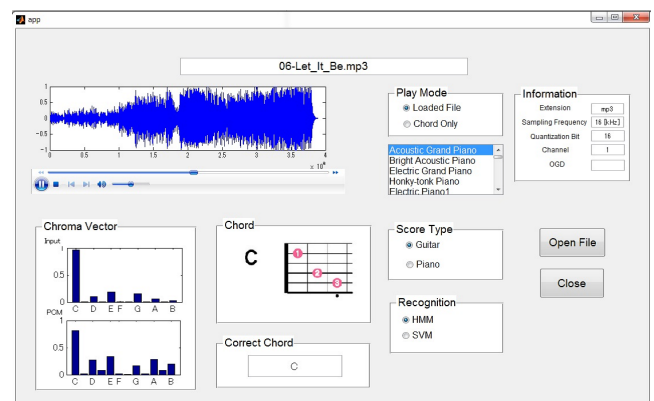


図 4 和音認識システム

参考文献

- [1] Supervised Chord Recognition for Music Audio in Matlab: <http://labrosa.ee.columbia.edu/projects/chords/>.

楽譜を用いた芸術的逸脱の統計的モデル化手法に基づく鍵盤楽器演奏の分析と再構築

奥村 健太, 酒向 慎司, 北村 正

演奏者による鍵盤楽器演奏の生成過程を捉え、演奏固有の特徴を楽譜との関係から説明する枠組みを紹介し、成果の展示を行う。図 5 に提案手法の概要を示す。

楽器演奏における演奏者固有の特徴は、演奏者の楽曲解釈に基づく芸術的な意図によって楽器の操作に生じた逸脱傾向に表現されると考える。本手法では、演奏とその楽譜の間で観測された任意の逸脱傾向に対して楽譜から得られる局所的な演奏指示を関連付けたモデルを個別に定義する。それらのモデルを楽譜の情報を制約として自動的に分類することにより、演奏者が生成しうる任意の逸脱傾向とその条件の組み合わせを木構造として体系的に表現できる。本手法による演奏分析の結果は楽器演奏に関する経験的な知見を裏付ける傾向を示しており、他のそのような知見の客観的な証明にも有用である [1]。

本手法の有用性については、既存演奏の分析以外の用途でも検証を進めている。その一例として、直近では未知の楽曲に自然な演奏表情を付与するための手法としての提案も行った [2]。提案手法では、既存演奏から選択された事例が持つ特徴の直接転写を行うことで、その演奏者に忠実な特徴を保持した仮想演奏を生成できる。任意の演奏指示に最適な演奏をあらゆる既存事例の組み合わせから探索することは現実的には困難であるが、モデルの自動分類により得られた木構造を活用することで、それを現実的なコストで実現できることを示した。

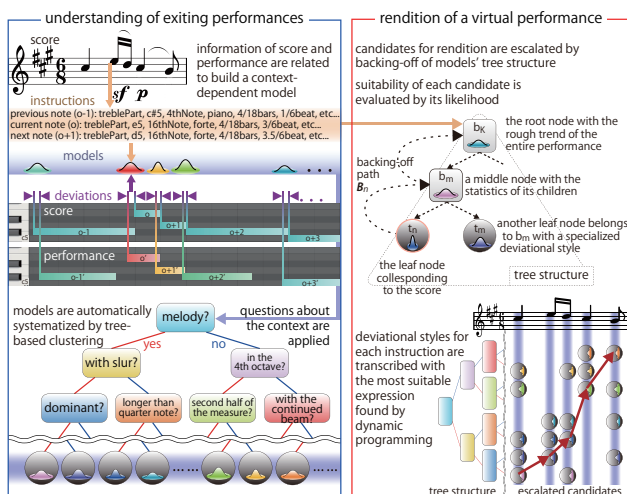


図 5 提案手法の概要

参考文献

- [1] 奥村 健太, 酒向 慎司, 北村 正: 楽譜に基づく鍵盤楽器の統計的モデル化手法, 情報処理学会論文誌, Vol.54, No.4, pp.1288-1301, 2013.
- [2] 奥村 健太, 酒向 慎司, 北村 正: 芸術的逸脱の生成モデルにおける決定木のバックオフに基づく鍵盤楽器演奏の自動表情付け, 情報処理学会研究報告, Vol. 2013-MUS-99, No.10, pp.1-6, 2013.

定常状態誘発反応の脳波可聴化システム

加庭 輝明

定常状態誘発反応の脳波 (EEG) 計測データを入力とする可聴化システムを紹介する [1]。定常状態誘発反応とは、短い間隔刺激を繰り返し提示することで、脳内で刺激に対するリズム同調が起こり、提示された刺激と同じ周波数で脳活動が現れる現象である。ブレインマシンインタフェースへの応用が検討されているが、安定した反応の検出には加算平均を取る必要があったり、発生と伝搬プロセスが明らかになっていなかったりする問題がある。我々はこれらの問題解決を目指し、データ観測時のモニタリングに向けた可聴化システムを構築した。

1. 可聴化の流れ

図 6 に可聴化の流れを示す。刺激周波数で誘発反応が発生する性質に着目し、短時間フーリエ変換を行い、刺激周波数に対応する周波数ピンのスペクトル情報を基に音合成を行っている。またその際に、正規化によって注目したいデータの性質を明確にすることで、ユーザーにとって認知的負荷の低い情報提示を行う。

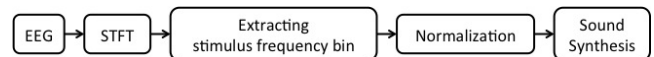


図 6 可聴化の流れ

2. 正規化・音合成の特徴

正規化は「チャンネル間の正規化」および「周波数ピン内の正規化」の 2 種類に分けられる。正規化の方法によって着目したいデータの性質を明確にすることがねらいである。正規化されたスペクトルの振幅値はパルス波の周期にマッピングされる。パルスの間隔の長短 (ピッチ) の変化から、誘発反応の発生や変化の様子を知覚することができる。

1. チャンネル間の正規化

全 EEG チャンネルにおける刺激周波数ピン中の最大振幅値を用いて正規化を行う。チャンネル間のパワーの比率が保存され、その情報が音に反映されるため、データの持つ空間的性質に着目できる。

2. 周波数ピン内の正規化

周波数ピン内の正規化では、それぞれの EEG チャンネルごとの刺激周波数ピン内の最大振幅値を用いて正規化を行う。周波数ピン内の平均パワーの大小に関わらず、信号の時間変化に着目できる。

謝辞 有益なコメントをいただいた美山千香士氏に感謝する。

参考文献

- [1] 加庭 輝明, 寺澤 洋子, 松原 正樹, Tomasz M. Rutkowski, 牧野 昭二: 脳活動の空間性と時間性に着目した EEG 可聴化, 日本音響学会 2013 年秋季研究発表会講演論文集, 3-2-4, 2013, (掲載予定).

体験！異なった時間軸を持つ複数の歌声をモーフィングしよう

河原 英紀, 森勢 将雅, 坂野 秀樹

これまで、速度の異なる2種類の表現で演奏された歌唱のレコーディングを、ライブでインタラクティブに操作してモーフィングする場合にも、ポストプロダクションで設定されたシーケンスに基づいてモーフィングする場合にも適用できる、時変多属性モーフィングを定式化・実装してきた [1, 2]。ここでは、新しいアイデアに基づいて拡張・定式化された、速度の異なる任意の個数の表現で演奏された歌唱レコーディングのモーフィング [3] を紹介する。この方法が提供する、多様な操作と表現の可能性を体験して頂きたい。

1. 時変多属性多事例間モーフィング

図7に3種類の表現をモーフィングしている際の状態表示を示す。左のカラーの表示は、モーフィング結果の時間周波数表現（周期性に起因する干渉の無いスペクトログラム）である。縦線は時間軸整合のための基準点、線上の印は周波数軸整合のための基準点を示す。右下は、3種類の表現のモーフィング率のベクトル時系列、右上は、それぞれの表現（細線）とモーフィング結果（太線）の基本周波数軌跡を示す。会場ではこれらの操作のための様々なGUIを用意する。

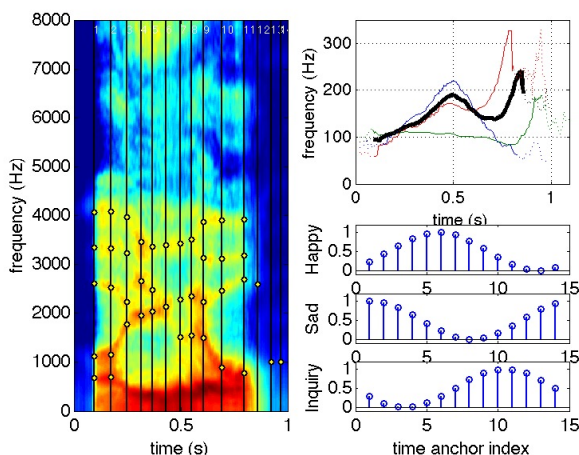


図7 Morphing status display

謝辞 本研究の一部は、科学研究費基盤(B)24300073および挑戦的萌芽 24650085 による。

参考文献

- [1] Hideki Kawaahra, Ryuichi Nisimura, Toshio Irino, Masanori Morise, Toru Takahashi, Hideki Banno, Temporally variable multi-aspect auditory morphing enabling extrapolation without objective and perceptual breakdown, Proc. ICASSP, Taipei, Taiwan, pp.19–24 (2009).
- [2] Masanori Morise, Masato Onishi, Hideki Kawahara and Haruhiro Katayose: v.morish '09: A Morphing-based Singing Design ICEC 2009, Paris, pp.185–190 (2009).
- [3] Hideki Kawahara, Masanori Morise, Hideki Banno, Verena G. Skuk, Temporally variable multi-aspect N-way morphing based on interference-free speech representations, (submitted)

タブレット型デバイスを用いた電子楽譜プラットフォーム「piaScore」

小池 宏幸

iPad上での電子楽譜の閲覧・配信のためのプラットフォーム piaScore[1] を紹介する。本プラットフォームは、「世界中の楽譜をあなたの手元に」をコンセプトに、演奏者なら誰もが感じる、重い楽譜の持ち運びや、見つからない楽譜をあれこれ探すことの苦勞など、楽譜に対するあらゆる悩みを解決するために開発されている。

1. 目標

より芸術的・娯楽的な演奏に向けたサポートや、新たな楽譜ビジネスの創出を目標とする。そのために、楽譜をデジタル化して持ち運べるようにするだけでなく、自動譜めくりや楽譜に関連した音楽・動画の視聴機能、ネットストアでの楽譜の購入・自作曲の販売など流通のデジタル化、日々の練習管理や練習法の自動サジェストなどの開発を進めている。

2. 実装した機能

演奏や練習に必要な機能にフォーカスした実装を行っている。具体的には、ジェスチャーによる譜めくり、デジタル書き込み、メトロノームやチューナーなどの楽器練習ツール、無料クラシック楽譜のダウンロード機能、楽譜購入機能などである。

3. 実際の用途

ピアノ、ギターなどの練習やコンサートでの利用を中心に、世界で多くのユーザに活用頂いており、2013年7月時点で累計20万ダウンロードを記録している。また、全日本ピアノ指導者協会や音楽大学の指導者には、授業やセミナーで活用頂いている。



図8 タブレット型デバイスを用いた電子楽譜プラットフォーム「piaScore」

参考文献

- [1] piaScore <http://piascore.com/>

拡張マルチスケールフラクタル次元を用いた環境音の類似検索システム

須之内 元洋, 田中 譲

我々が開発した拡張マルチスケールフラクタル次元という音響特徴量を用いた、環境音を対象としたコンテンツベースの類似検索システム Shirakaba を紹介する。サウンドデザインや楽曲制作の支援等を目的とした環境音の類似検索を実現するためには、普段耳慣れない音や多様な音源が混在する環境音の特徴を、上手に記号表現できる特徴量抽出が大きな課題となる。

1. 拡張マルチスケールフラクタル次元

音響信号の波形を形態的にとらえることにより波形のフラクタル次元を算出できる。フラクタル次元を算出する際に利用する円の半径をある範囲で変化させながら、複数のフラクタル次元の値を算出することで、マルチスケールフラクタル次元 (MFD) という音響特徴量が従来提案されている。我々は、様々な環境音の MFD の性質を調査したうえで、MFD を算出する際に利用する円の半径の変化範囲を広範囲に拡張し、対数軸で円の半径を変化させることとし、さらに得られた MFD の時間方向の分布を表すヒストグラムを作成して、拡張マルチスケールフラクタル次元 (EMFD) という音響特徴量として定義した。代表的な音響特徴量である MFCC にはない特徴が EMFD に含まれていることを確認し、EMFD の有効性を示した [1]。

2. 類似検索システム Shirakaba

図 9 は環境音の類似検索システム Shirakaba の検索結果画面である。Freesound (<http://www.freesound.org/>) からインポートした 3,000 件の環境音が収められており、任意の音源を指定して検索を実行すると、類似音源の候補を類似度が近い順に列挙する。

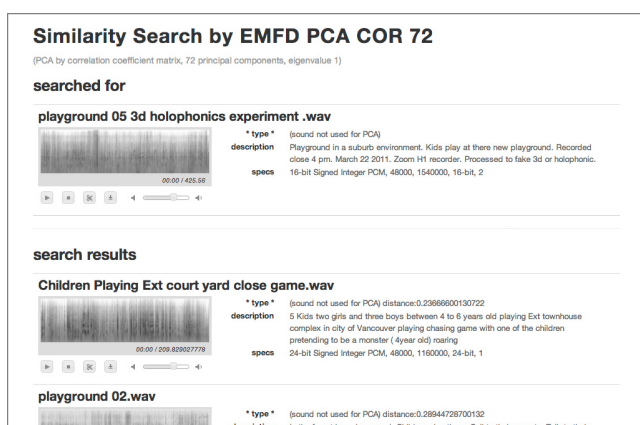


図 9 Shirakaba の類似検索結果画面

謝辞 有益なコメントを頂いた産業技術総合研究所の後藤真孝博士に感謝する。

参考文献

- [1] M.Sunouchi and Y.Tanaka, "Similarity search of freesound environmental sound based on their enhanced multiscale fractal dimension," in Proc. SMC2013, Stockholm, Sweden, Aug.2013, pp.715-721.

演奏用インタフェースのプロトタイピングのための PC 用キーボードカバー

竹川 佳成, 寺田 努

現在、PC 用キーボードは入力速度と入力精度の高さからコンピュータの入力インタフェースとして不可欠である。また、PC 上で動作する作編曲アプリケーションや演奏アプリケーションが多数提案されており、Max/MSP や Processing などのプログラミング言語を利用して手軽に自作の音楽アプリケーションを制作できる環境が整ってきた。アプリケーションを操作するための入力インタフェースを手軽に制作できれば直観性や操作性の向上が期待できる。そこで、我々は、PC 用キーボードを単純なボタンの配列とみなし、PC 用キーボード上に設置することで新たな入力インタフェースを作成できるキーボードカバーを提案している [1]。

提案するキーボードカバーは薄いプラスチック板や厚紙からできており、これに任意形状の切り込みをいれ突起を付与することで、操作したいアプリケーションに特化したキー配列をもつ入力インタフェースを制作できる。例えば、図 10-(a) に示すようにプラスチック板に鍵盤を模した切り込みをいれ、キーボードカバー上のキーに特定の PC 用キーボード上のキーを押すための突起を貼り付ける。これによりキーボード上に鍵盤楽器のレイアウトをもつキーボードカバーを配置できる。図 10-(b) に示すように異なる高さの突起を鍵の下に 2 箇所配置することで、各キーボードキーが押されるタイミングが変化し、打鍵の強さを検出できる。このキーボードカバーは、学校でパソコンを使いながら作曲理論や楽器の構造など音楽学の講義をする場合の教材として手軽に使用できる。デモセッションでは、上記以外にさまざまな演奏用キーボードカバーを展示する予定である。提案するキーボードカバーは突起の位置や高さを工夫することで、キーの重さなどの打鍵感を変えて、操作性を含めて柔軟にカスタマイズすることもできる。実際に提案するキーボードカバーを体験していただくことで、キーボードカバーの有用性や応用について議論したい。



図 10 鍵盤型キーボードカバーの外観

参考文献

- [1] 竹川佳成, 寺田努: PC 用キーボードのキー配列を柔軟に変更可能なキーボードカバーの提案, エンタテインメントコンピューティング 2012, pp. 56-65, 2012.

MIDI フィッティングシステム

土屋 政人, 高宗 典玄, 亀岡 弘和,
堀内 靖雄, 橋田 光代, 嵯峨山 茂樹

本システムは楽譜情報を用いて音響信号を再現するような MIDI パラメータを推定する。音響信号から MIDI を得ることができれば、音響信号データベースを MIDI 化するだけでなく、演奏表情付けの学習データとして用いることができ、奏者の特徴抽出・分析・転写などにも応用することができる。関連研究として音響信号と記号・MIDI レベルとのアラインメントを推定するものが存在するが、我々は観測信号を音響的に再現する MIDI パラメータを推定するところに主眼を置いている。本システムは一音ごとのスペクトログラムテンプレートを使って MIDI パラメータを推定する自動フィッティングと人間の視聴覚によって推定を行う手動フィッティングの2つのサブシステムから成る。

1. 自動フィッティングサブシステム

このサブシステムでは観測音響信号に近いスペクトログラムを生成するような MIDI パラメータを推定する。予め作っておいた単音スペクトログラムテンプレートを用いて、補助関数法によって観測信号スペクトログラムとの距離が最も小さくなるような局所解を探索する。しかし、この手法は初期値依存性が強く、多くの局所解が存在するため、前処理として楽譜情報を等価変換した“機械的演奏”の MIDI と観測音響信号との大まかな時間アラインメントを取る必要がある。そこで音色の違いに対して誤り精度を低減できる、Ellis らによる手法を用いた。詳細は [1] を参照されたい。

2. 手動フィッティングサブシステム

手動フィッティングでは図 1 のようにスペクトログラム上に各ノート半透明で重畳表示し、人間の視聴覚情報に基づく判断を基本とした MIDI パラメータの推定を行う。これによって自動フィッティングにおけるエラーを修正することができる。

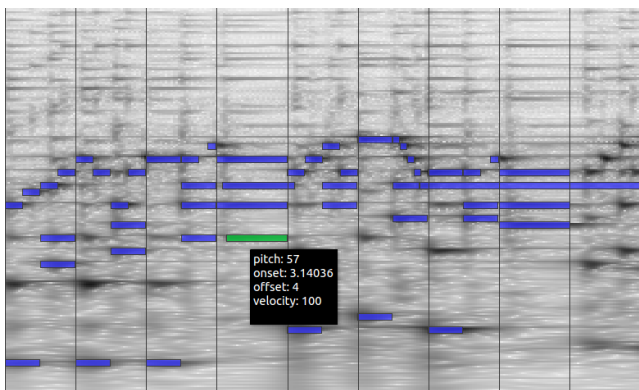


図 11 開発中の GUI

参考文献

- [1] 高宗他, 音講論, “補助関数法による音楽音響信号への MIDI 信号の詳細なフィッティング”, Sep. 2012.

音楽音響信号の曲調変換システムに向けて

中村 友彦, 亀岡 弘和,
中村 栄大, 小野 順貴, 嵯峨山 茂樹

楽曲の音響信号を入力として、音階やリズムなどを適切に変換することにより、曲調を変換可能なシステムの実現に向けて議論する。曲調は音階やテンポ、リズムなど様々な音楽の構成要素に起因しており、ユーザが楽曲を所望の曲調にするため作曲や編曲を行う際に、いかにこれらの要素を変換するかが問題となる。そのため、音楽の構成要素の変換を自動化すればこのようなユーザを補助・支援できる。そこで、本研究では音楽信号処理と音楽理論を併用し、音楽の構成要素を変換するアルゴリズムの開発を目指す。

1. 曲調変換のための課題

曲調変換のためには、以下の2つの課題がある。(a) 音響信号処理の課題として、所望の音高やテンポへの音楽音響信号の加工である。これには、ノート単位で音響信号を編集可能な Music Factorizer [1] をベースとし、実用に耐える程度の音質を保持できるようにアルゴリズムを改良する予定である。(b) 音楽記号処理の課題は、音楽の構成要素をいかに変換するかである。これには、確率的な枠組みを用いて既存の楽曲から学習することを考えている。

2. 曲調変換システムの開発状況

現在、Music Factorizer [1] をベースとし、音階の変換およびリズムの変換システムを開発中である(図 12)。長音階と旋律的短音階を対象として、主音からの度数を保持する変換則に従い、単旋律の音響信号を変換できる。また、D.P.W. Ellis らの手法 [2] を用いて入力音響信号から推定した拍位置を元に、1小節を単位として音響信号を加工することにより、リズムの変換を行おうとしている。

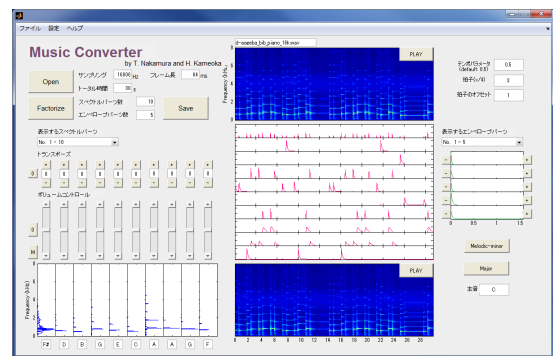


図 12 開発中の曲調変換システム

参考文献

- [1] 亀岡 弘和, Jonathan LeRoux, 大石 康智, 柏野 邦夫: Music Factorizer: 音楽音響信号をノート単位で編集できるインタフェース, 音情研報, 2009-MUS-81-9, 2009.
- [2] D.P.W. Ellis, “Beat Tracking by Dynamic Programming,” J. New Music Research, vol.36, no. 1, pp.51-60, Mar, 2007.

Eurydice: 任意の弾き直し・弾き飛ばしにも追従する自動伴奏システム

中村 友彦, 中村 栄太,
齋藤 康之, 小野 順貴, 嵯峨山 茂樹

Eurydice は人間の MIDI 楽器演奏に合わせて、楽譜に基づき自動的に伴奏を行うシステムである。音高間違い、音抜け、余分な音の付加、遠方への弾き飛ばし、弾き直しを含む演奏を隠れマルコフモデルによりモデル化し、高速 Viterbi アルゴリズムを用いて、実時間で多様な演奏に瞬時に追従する(文献 [1,2])。また、テンポや強弱の変化などの奏者の演奏表現をリアルタイムで伴奏に反映する。練習支援としてピアノの左手パートや連弾の伴奏を担当するだけでなく、ピアノ協奏曲でのオーケストラ伴奏を担当することも可能である。

1. Eurydice の仕様

Eurydice は、MIDI 演奏信号を入力として、演奏楽譜位置、テンポ、強弱などの情報を得る楽譜追跡部分と、演奏に同期した伴奏の MIDI 信号を出力する伴奏再生部分からなっている。ユーザは演奏する楽曲を MIDI ファイルとして与え、GUI を通じて、演奏する声部と自動伴奏を行う声部を指定できる。アプリケーションは Windows, Mac OS X などの環境で動作する(図 13)。

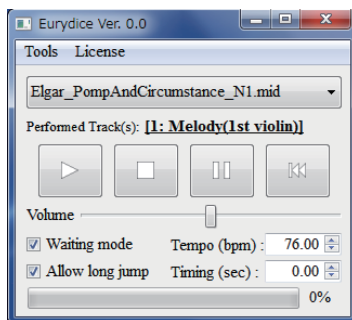


図 13 Eurydice ver. 0.0

2. Eurydice 0.0 の公開予定

我々は、Eurydice を非商用使用に限り <http://hil.t.u-tokyo.ac.jp/software/Eurydice/> で無償配布を予定している。配布予定の Eurydice ver. 0.0 は文献 [2] を実装したものであり、任意箇所から任意箇所への弾き直しや弾き飛ばし、演奏誤りを含む演奏でも実用に耐える速さで自動伴奏を行える。

謝辞 このソフトウェアの開発は、科学研究費補助金・基盤研究 B (課題番号 17300054)、基盤研究 A (課題番号 23240021)、および科学技術振興機構 CREST プロジェクトの補助を受けて行いました。また、開発初期には、戦略ソフトウェア創造人材養成プログラムにおいて、平木 敬教授、稲葉 真理 准教授、金子 勇氏、土村 展之氏の各氏から有益な助言を頂きました

参考文献

- [1] 武田 晴登, 西本 卓也, 嵯峨山 茂樹: “HMM による MIDI 演奏の楽譜追跡と自動伴奏,” 情報処理学会研究報告, vol.2006-MUS-66, pp.109-116, Aug. 2006.
- [2] 中村 栄太, 武田 晴登, 山本 龍一, 齋藤 康之, 酒向 慎司, 嵯峨山 茂樹: “任意箇所への弾き直し・弾き飛ばしを含む演奏に追従可能な楽譜追跡と自動伴奏,” 情報処理学会論文誌, vol. 54, no. 4, pp.1338-1349, Apr. 2013.

複数の実ユーザとマルチエージェントによる協調演奏システム

中村 亮介, 渡辺 和泉,
後藤 敏行, 田村 直良, 島田 広

筆者らは、実ユーザと計算機内部の複数のエージェントが楽曲を共演するシステム [1] の開発を進めている。本デモセッションでは、MIDI 鍵盤を用いてユーザと複数の演奏エージェントが 1 つの楽曲を演奏する。

1. 協調演奏システムの概要

協調演奏システムは各パートを担当する複数のエージェントで構成されている(図 14)。各エージェントは独立した MIDI シーケンサを備え、リアルタイムに共有メモリを介して演奏情報を交換しながら、演奏モデルに基づいてそれぞれが判断して、ピッチ、ペロシティ、テンポなどを変化させながら演奏を行うことができる。そして、エージェントをユーザに置き換えることによって、MIDI 楽器を用いてユーザが演奏に参加することができる。また、エージェントを表すグラフィックスの提示など、視覚的な出力も備えている。

2. 演奏モデルと最適化

ユーザの演奏位置推定には、MIDI ベースの楽譜と演奏との音列マッチングを利用している。この際、トリルやフェルマータ等の奏法記号は、演奏時間や回数に自由度があるため、演奏位置の推定が困難になるという問題がある。これに対して本研究では、奏法記号をメタイベントと呼ばれる拡張可能なメッセージとして MIDI シーケンスに埋め込むことで対応している。

さらに、それぞれエージェント間の演奏パラメータの影響は一律ではなく、演奏者同士の空間的配置を考慮した「影響度」という概念を導入してモデル化している。また、複数の実ユーザが演奏に参加する場合には、単純に楽譜追跡結果を利用すると伴奏が乱れる恐れがある。本システムでは演奏者の「信頼度」という設定を提案し、この問題の解決を図る。

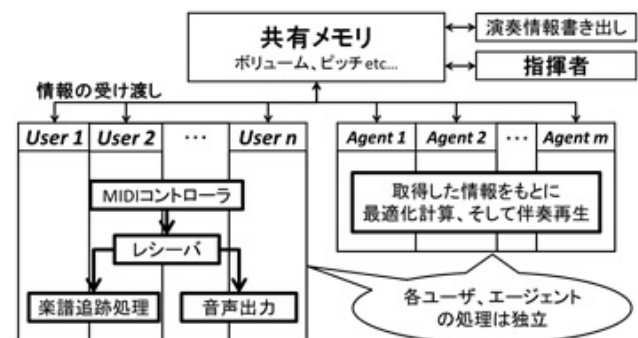


図 14 システム概要

参考文献

- [1] 石塚, 中村, 後藤, 田村, 島田: “仮想空間オーケストラにおけるマルチエージェント協調演奏シミュレーション”, 映像学技報 ME, Vol.37, No.7, pp.99-102, 2013.

Songrium: コンテンツのつながりを楽しむ音楽視聴支援サービス

濱崎 雅弘, 後藤 真孝, 中野 倫靖

音楽コンテンツ間のつながりを意識しながら, Web 上で新たな音楽コンテンツに出会うことができる音楽視聴支援サービス「Songrium」(ソングリウム) (<http://songrium.jp>)を開発した。

1. つながりを生かした音楽視聴支援

音楽と音楽の間にはつながり(関係)がある。類似関係や派生関係, クリエータの人間関係などはその一例であり, ある楽曲は他の楽曲と様々な関係を持っている。関係が, 新たな楽曲に出会うときの手がかりになることも多い。しかしこうした様々な関係に基づいた音楽コンテンツとの出会いは十分に支援されていなかった。

そこで我々は, 音楽の多様な関係性に着目した音楽視聴支援サービス Songrium を開発し, 2012 年 8 月に 版を公開した [1-3]。Songrium は自動的に抽出・推定した派生関係や類似関係を可視化し, さらに二つの音楽コンテンツ間に「矢印タグ」を定義して共有するソーシャルアノテーションを可能にする Web サービスである。

2. Songrium

図 15 に Songrium の画面例を示す。Songrium はニコニコ動画 (<http://nicovideo.jp>) に投稿された歌声合成関連動画を対象としている。「惑星ビュー」は自動抽出した楽曲と派生作品との関係を可視化し, 派生作品群の俯瞰だけでなく, 人気作や最新作を容易に見つけられる。「矢印タグ」は, 音楽コンテンツ間の明示的あるいは暗黙的な関係に対するソーシャルタギングを可能にする新たな枠組みであり, 人々が矢印タグを自由に名付けて定義し, 共有していくことで, 様々な関係を扱えるようになっていく。「音楽星図」は音響特徴量に基づき楽曲を平面配置したもので, 約 10 万曲を一望できる。



図 15 Songrium の画面例

謝辞 Songrium の Web サービスの実装を担当して頂いた石田 啓介 氏に感謝する。本研究の一部は JST CREST 「OngaCREST プロジェクト」の支援を受けた。

参考文献

- [1] 濱崎雅弘, 後藤真孝: Songrium: 多様な関係性に基づく音楽視聴支援サービス, SIGMUS, Vol.2012-MUS-96, 2012.
- [2] 濱崎 雅弘, 後藤 真孝: Songrium: 関係性に基づいて音楽星図を渡り歩く音楽視聴支援サービス, WISS 2012.
- [3] M. Hamasaki, M. Goto: Songrium: A Music Browsing Assistance Service Based on Visualization of Massive Open Collaboration Within Music Content Creation Community, ACM WikiSym+OpenSym 2013.

和声学に基づく合唱譜自動作成システム

Benjamin Evans, 棟方 渚, 小野 哲雄

与えられたバスパートの楽譜に, 和声学の規則に従ってソプラノ・アルト・テナーの 3 パートを記譜し合唱譜を作るシステム「CMY」を紹介する。既存の作曲システムの多くは, 作曲過程が確率モデルやパターンマッチングなどを利用し, ユーザによる作曲プロセスの音楽学的変更は難しい。CMY は入出力に楽譜データを利用し, 作曲規則も全て和声学に従って表現されている。そのため, 出力される楽曲のみならず作楽過程そのものも, ユーザによって調整しやすくなる。

1. CMY システム

図 1 に CMY の実行画面の例を示す。CMY はバスパートの楽譜 (MusicXML 形式データ) を読み込み, ユーザが指定する和声学などの規則を利用して残り 3 パートを配置し, その結果を出力する。一般に, あるバス譜に対して複数の作曲の可能性があるため, CMY はユーザが指定する評価基準などを利用し, 評価値の高い楽曲を出力する。

2. CMY で利用する和声規則

複数の音を同時に発するときの音の協和・不協和とその連結に関する規則などを体系的に述べた和声学は, 多くの作曲活動の基礎となっている。ここでは日本の音楽教育の基準 [1] となっている「芸大和声」[2] に焦点を絞り, その中から基本的な諸規則を十数個実装した。ユーザは作曲に利用する(しない)規則を選択する。また, 和声規則にその適用度を指定し, 時折和声学に反するような楽曲も出力することができる。CMY を利用することにより, ユーザは音楽規則を任意に選択し, 新しい響きを持つ楽曲の作曲に挑戦できる。

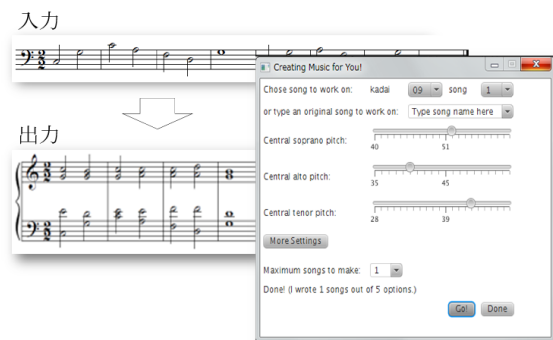


図 16 CMY 概要

参考文献

- [1] 三浦 雅展, 江村 伯夫 (2012). “和声理論に基づいた作・編楽曲システム”, システム/制御/情報, 56 巻 5 号, pp.213-218
- [2] 池内 友次郎, 島岡 譲ら (1964). 『和声 理論と実習』 I

自由視点映像と聴覚提示に基づく スマートフォン向けインタラクティブ ミュージックビデオアプリケーション

堀内 俊治, 三功 浩嗣, 加藤 恒夫, 内藤 整

本アプリケーション [1] は, スタンドアローンのスマートフォン上で単体動作する. オーディオレンダリングは, 頭部伝達関数によるバイノーラル再生に基づき実装した. ビデオレンダリングは, 演者などのオブジェクト毎に, ビルボードと呼ばれる簡易 3 次元モデルを構築する方式に基づき実装した. オーディオソースおよびビデオソースから抽出したテクスチャはアプリケーションに埋め込んでいる. 本アプリケーションは, 音楽レーベルの公式コンテンツとして, Android と iOS 向けにリリースされた.

本アプリケーションでは, 視点は, タッチスクリーン上における以下のジェスチャにより, 操作することができる: 1) スライドレフト・ライトにより, ステージの周りを回る, 2) ピンチイン・アウトにより, 遠ざかる・近づく. 更に, 3) タッチ&ホールドおよびドラッグにより演者の位置を変更することができる. 加えて, シーンの変更により, 例えば, ライブハウスやコンサートホールの音響効果に切り替えられる. オーディオソースは, 演者ごとに個別のマイクロホンにより録音した. ビデオソースは, 複数のカメラにより撮影した. 演者ごとのテクスチャは, ビデオソースより抽出した. オーディオソースと抽出したテクスチャは, それぞれ AAC ADTS および H.264 ES フォーマットで保存した. 図 17 は, スクリーンショットである. 本デモンストレーションでは, オブジェクト (演者) 数は 4, カメラは 6 台 (円状に 60 度ごと), 最大フレームレートは 10 fps, およびサンプリングレートは 48 kHz とした. 1~3 Mbps 程度のビットレートである. 30 s のコンテンツの場合, アプリケーションサイズは 20 MB 以下となる.

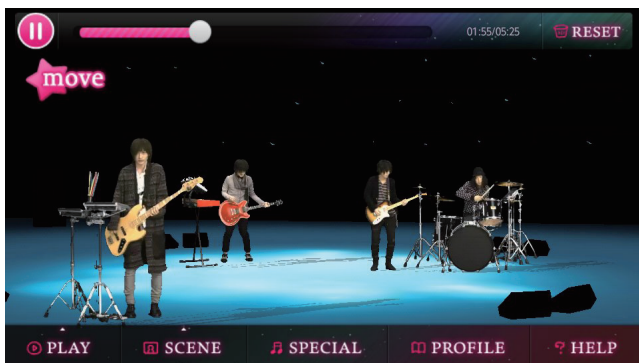


図 17 スクリーンショット

謝辞 ご協力いただいた Applicat Spectra および株式会社 A-Sketch に感謝する.

参考文献

- [1] Horiuchi, T., Sankoh, H., Kato, T., and Naito, S.: Interactive Music Video Application for Smartphones Based on Free-viewpoint Video and Audio Rendering, Proc. 20th ACM International Conference on Multimedia (ACM Multimedia 2012), pp.1293-1294 (2012).

リハビリ応用に向けた 筋電可聴化フィードバックシステム

松原 正樹, 井口 正樹, 門根 秀樹, 寺澤 洋子, 鈴木 健嗣

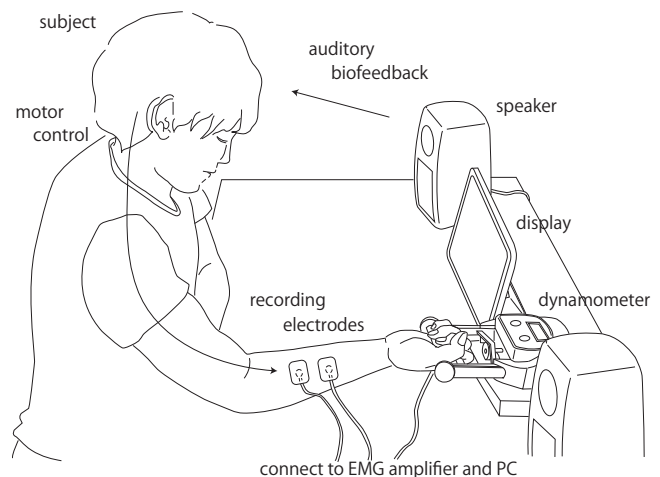


図 18 システム実験時のセットアップ

リハビリ応用に向けた筋電可聴化フィードバックシステムを紹介する [1]. 内的な生体信号を視覚や聴覚を通して知覚できるようにフィードバックすること (バイオフィードバック) は, 患者自身の運動感覚と実際の動きを関連づけて認識できるため, 運動療法を用いたリハビリにおいて有効であると考えられる. 特に可聴化によるフィードバックは視覚障害を持つ人に有効なだけでなく, 移動を伴う運動時 (歩行訓練) や視線が固定されている時 (ベッドサイド訓練) にも有効である. 本研究は上肢の筋活動電位を可聴化フィードバックすることで腕や手の適切な運動を学習できるようなシステムの設計を目的とする.

図 18 にシステム実験時のセットアップの概略を示す [1]. 生体アンプを用いて筋電を取得し, 入力パラメタとして音の振幅に変換し可聴化を行った. 実験では全盲 10 名と健常者 11 名に対して手を任意の強さで握る運動における可聴化フィードバックの有効性の検証を行い, 可聴化により適切な運動が行えるようになったことが確認された. 可視化フィードバックとの比較実験も行い, 特に全盲被験者においては可聴化による主観的理解度が健常者の可視化の結果と同等であることが示された.

本デモセッションではこれまでに提案した可聴化手法 [2] をもとにシステムを構築し, より自由な運動による可聴化フィードバックを体験してもらい. 複数の可聴化フィードバック体験を通じて比較し議論を行う.

謝辞 有益なコメントをいただいた美山千香士氏に感謝する.

参考文献

- [1] M. Iguchi, M. Matsubara, H. Kadone, H. Terasawa and K. Suzuki: "How is Auditory EMG Biofeedback effective for Blind People?," International Conference on Auditory Display 2013, pp. 307-310, 2013.
- [2] M. Matsubara, H. Terasawa, H. Kadone, K. Suzuki and S. Makino: "Sonification of Muscular Activity in Human Movements using the Temporal Pattern in EMG," Asia Pacific Signal and Information Processing, No. OS.6-BioSPS.1-2, pp.1-5(CD-ROM), 2012.

音楽理論 GTTM を用いた議論タイムスパン木生成システムのプロトタイピング

三浦 寛也, 森 理美

1. はじめに

ディスカッションマイニング (DM)[1] とは、会議における活動を複数メディアで記録し、そこから再利用可能な知識を抽出するための技術である。我々は、音の時系列イベントを構文解析する技術である音楽理論 GTTM[2] の分析手順に基づき、会議記録の発言の重要度を階層的に表現する議論タイムスパン木 (図 19) を提案した [3]。本セッションでは、現在プロトタイピングしている議論タイムスパン木の生成システムを紹介する。

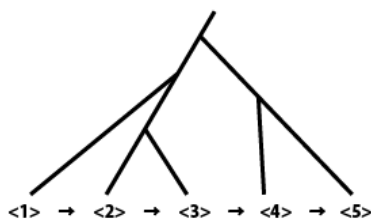


図 19 議論タイムスパン木

2. 議論タイムスパン木の生成手順

本システムの分析対象は DM システムでの会議データである。DM 木とは先行する発言が有るという関係に基づいて構成された木構造である。図 1 は、発言番号 <1> から <2> が生じ、さらなる発言が生じたことを表している。我々は以下 2 段階の処理によって、DM 木を基に議論タイムスパン木を生成する。(1) グルーピング獲得: DM 木に含まれるグループ (ゲシュタルト) の発見。(2) 重要発言の選定: あるグループ全体の時間幅 (タイムスパン) を代表する重要発言の選定。図 19 の議論タイムスパン木は各発言の重要度の順序が高い方から順に <2>, <5>, <1>, <4>, <3> を表現している。

3. 実装

議論タイムスパン木生成に関するルール実行管理として、選好ルールの競合解消と階層構造の獲得が主要な問題である。これらの問題に対処するため、ルールを評価値として導入し、以下 2 項目を判定する。(1) 発言間に生じる境界、(2) 発言が内包する情報量の大きさ。その結果から、相対的なグループ発見をもとにボトムアップに重要発言の選定をおこない、議論タイムスパン木を生成していく。

参考文献

- [1] 長尾研究室: ディスカッションマイニングプロジェクト, <http://dm.nagao.nuie.nagoya-u.ac.jp/>
- [2] Lerdahl, F. and Jackendoff, R.: A Generative Theory of Tonal Music, The MIT Press (1983).
- [3] 三浦寛也, 富樫健太, 浜中雅俊, 長尾確, 東条敏, 平田圭二, 音楽理論を応用したディスカッションマイニングにおけるタイムスパン木と延長木の自動生成について, 情報処理学会デジタルコンテンツクリエーション研究会, 2013-DCC-3, No.9 (2013).

DIPS5 for Max

葉 孝之, 松田 周

DIPS は 1997 年に松田が jMax のためのプラグイン・ソフトウェアとして開発を始めたリアルタイム映像処理を実現する Max のオブジェクト群です。ベルリンで開催された ICMC 2000 において発表されました。2006 年には Mac 上の Max/MSP で稼働する DIPS3, 2007 年にはよりユーザーフレンドリーなプログラミング環境とライブラリが追加された DIPS4 がリリースされました。この夏リリースの DIPS5 では最新の映像技術や周辺機器への対応がアップデートされ、チュートリアル・パッチ、サンプル・パッチの充実が図られています。

1. DIPS の特徴

当初、Max のパッチ内で OpenGL プログラミングを実現することを目的に開発された DIPS ですが、DIPS4 からは OpenGL を知らない作曲家やアーティストでも容易に映像処理プログラムを組み、音響処理との連携を実現することができるように DLib と Dfx というオブジェクト群が用意されました。Dlib はプログラミングを簡素化するためのユーティリティ・オブジェクト群であり、Dfx は OpenGL コンテキストベースで動作するビデオ・エフェクター集です。更に、OpenGL やシェーダのプログラミング、モーション・ディテクトやフェース・トラッキング、オプティカルフロー等の映像分析機能、Quartz Composer など様々なメディアアファイルへのインターフェースが充実され、HD 以上の解像度での映像生成も現実的になりました。

現在、DIPS は 300 以上のオブジェクトで構成されており、ランカスター大学や国立音楽大学等で Max 学習クラス、インタラクティブ・マルチメディア・アート創作実習クラスにおいて活用されています。学生は初めての授業からビデオ入力にさまざまなエフェクトをかけることができるようになり、後には Max の Application 化を利用することにより VJ ソフトを制作することも可能です。プログラミングの容易さは創作活動ばかりでなく、教育現場においても威力を発揮しています。

DIPS は葉, 松田のほか、高橋深太郎, 森本洋太, 福田拓人等によって開発が続けられています。

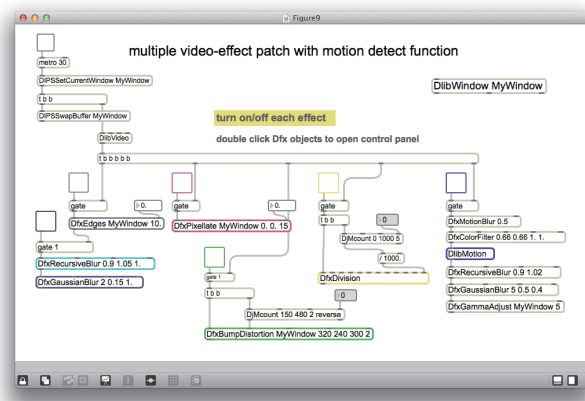


図 20 DIPS のビデオエフェクター・パッチ