

2 群(画像・音・言語) 9 編(音楽情報処理) 3 章(スタイル・ジャンル)

3-1 スタイル・ジャンル

(執筆: 平田圭二)

3-1-1 スタイル・ジャンルの意味

多数の人が所属するコミュニティに多数の楽曲が存在するとスタイルやジャンルが発生する。この時、1 曲だけではスタイルやジャンルとは呼ばれないとすると、同じような楽曲が何曲くらい集まったらスタイルやジャンルと呼ばれるようになるのだろうか。また、スタイルやジャンルは一般に階層構造を成しているが*、それは、あまりに幅の広い楽曲群を 1 つのスタイルやジャンルにくくるのは不自然だからである。では、スタイルやジャンルを定めている境界線のある曲が越えているかどうかはどのようにして決まるのだろうか。つまり、スタイル・ジャンルは、個々の楽曲間に共通する要素が人々に認識されるようになることと(前者の共通性の認識)、個々の楽曲を比べた時お互いどこを特異な部分と認識されるようになること(後者の特異性の認識)の 2 面から決まる。

音楽辞典¹⁾によれば、音楽学におけるスタイル(様式)とは、音楽作品の持つ具体的な特性に関して形式的特徴として現れる類型のこととある。類型の例としては、楽式(楽曲の形式)、表現材料や題材(神話、事件)、楽器(ピアノ、ヴァイオリン)、時代(バロック、現代)、個人の素質、人格、技巧(バッハ、モーツァルト)、民族(スペイン、中華、アフリカ)、流派(ウィーン、ヴェネチア)等が挙げられる。ジャンルも同様に、何らかの音楽的な分類を意味する。これらの類型は、共通性の認識と特異性の認識が互いに影響を及ぼしあいながら、時間をかけてコミュニティ内で到達した合意の一部であり、共通性の認識と特異性の認識のトレードオフと言ってよいだろう²⁾。

しかし、音楽情報処理におけるスタイル・ジャンルの意味は、音楽用語のそれとは大きく異なっている³⁾。音楽情報処理におけるスタイル・ジャンルとは、楽曲群を機械的に分類するために決められた名前(ジャンル名、スタイル名[†])の集まりのことである。音楽用語としてのスタイル・ジャンルのように音楽的な根拠をもってその名前が付けられる場合もあれば、音楽的な理由とは無関係な理由でクラスタリングを行い名前が付けられる場合もある(例えばビジュアル系など)。

3-1-2 2 つのアプローチ

自動スタイル・ジャンル分類のアプローチは大きく、音響信号に基づく方式(content-based)とメタデータに基づく方式(metadata-based)に分けられる。これは、楽曲間の類似度あるいはアーティスト間の類似度を算出するのに、音響信号の音響特徴量を使う方法とメタデータを使う方法が有効だからである。前者の音響信号に基づく自動スタイル・ジャンル分類に関する優れたサーベイとして文献⁴⁾がある。後者には、様々な手段で獲得されたメタデータに基づく方法⁵⁾や、Web 上での共起単語を用いた協調的フィルタリングに基づく方法⁶⁾等がある。

* 場合によっては、1 つの楽曲が複数のジャンルに属することもあり、厳密には階層構造ではない。

† ジャンル名の例にはジャズ、ポップ、ロック、クラシック等がある。スタイル名は一般に、ジャンルよりさらに細分化された分類を指す場合が多い。例えば、ロックはさらにハードロック、ヘビメタル、プログレ、グラム、パンクなどサブカテゴリに分かれている。

一般に、楽曲群をいくつかのクラスに分割する属性や尺度は、上述の共通性の認識や特異性の認識に対応して、音響的なものからテキストによるものまで多種類考えられ、それらは通常異なるクラスタリング結果をもたらす。例えば和声進行という尺度による楽曲群の分割と、テンポという尺度による楽曲群の分割は異なる。自動スタイル・ジャンル分類の研究とは、互いに異なるクラスタリング結果をもたらすような尺度を適切に組み合わせて、目標とするスタイル・ジャンル分類を精度高く実現する課題と換言できる。さらに、付与されるスタイル・ジャンル名が2個以上の場合もある。

音楽の専門知識のない人でもある程度安定してスタイル・ジャンル分類できるということから、それほど高度な音楽知識を仮定せず、人が普通に持っているような識別能力を実現すれば、ある程度まで自動的にスタイル・ジャンル分類できるという仮説が支持されている。実際、音響信号から様々な音響特徴量を取り出して特徴ベクトル間の距離を算出するアプローチでも用途を限定すれば実用的に十分な場合がある。しかし、人でも識別の間違いが起きるようなスタイル、例えば時代、民族、流派を音響特徴量だけで識別するのは難しい。このような場合は、上述の2つのアプローチを混合したり⁷⁾、高度な音楽知識を利用することが効果的である。

3-1-3 研究事例

スタイル・ジャンル分類に利用される音響的特徴量の多くは、人が楽曲を聴取する時に注意を向けている特徴や部分はどこかという内省に基づいて見いだされたものである。音楽の3要素は旋律、和声、リズムと言われるが、この内、旋律と和声に関する特徴はスペクトログラムのテクスチャとして出現する音響的特徴に関連付けられ、リズムに関する特徴は時間方向の周期的パターンとテンポに関連付けられることが多い。着目した特徴に関する情報を楽曲全体に渡って収集する方式 (bag-of-frames) と、時間順序に沿って収集する方式がある。音色、テンポ、旋律の各マクロ情報を用いた研究の最初期のものに⁸⁾がある。ミクロ情報として、リズム⁹⁾ 低音旋律¹⁰⁾、和声認識^{11, 12)} などを用いる研究がある。和声認識をしたあと、時間軸に沿って和音進行どうしの類似度を比較するのに Dynamic Programming (DP) や Hidden Markov Model (HMM) が用いられることが多い。次に、所与のスタイル・ジャンル名の集合と特徴量ベクトルの集合を対応付けるため、パターン認識を行ったり、統計的なモデルを仮定してそのパラメータを機械学習する⁴⁾。一方、記号処理レベルで、和声に基づくジャンル分類を行う研究もある¹³⁾。

スタイル・ジャンル分類に利用される重要な音響的特徴量や記号的特徴量に関しては、毎年 MIREX (Music Information Retrieval Evaluation eXchange) というコンテストが開催されている¹⁴⁾。音楽情報検索 (MIR) は記号処理、信号処理、テキスト処理など多分野にまたがる研究領域なので、最新の MIR アルゴリズムや MIR システムの性能比較は容易ではない。そこで、Audio Genre Classification, Audio Artist Identification, Symbolic Melodic Similarity などの要素技術に細分化し、与えられたベンチマークに関して性能を直接比較する。第1回めの MIREX コンテストは ISMIR 2005 (London) 会期中に開催された。

参考文献

- 1) 新音楽辞典 楽語, 音楽之友社 (1977).
- 2) Cory McKay and Ichiro Fujinaga, Musical genre classification: Is it worth pursuing and how can it be improved?, in Proc. of ISMIR 2006.

- 3) Jean-Julien Aucouturier and Francois Pachet, Representing Musical Genre: A State of the Art, *Journal of New Music Research*, Vol.32, No.1, pp.83-93 (2003).
- 4) George Tzanetakis, 音響ベースの音楽信号分類, *情報処理* Vol.50, No.8, pp.746-750, 角尾衣未留 (訳) (2008).
- 5) Dan Ellis, Brian Whitman, Adam Berenzweig, and Steve Lawrence, The quest for ground truth in musical artist similarity, in *Proc. of ISMIR 2002*.
- 6) Peter Knees, Elias Pampalk, Gerhard Widmer, Artist Classification With Web-Based Data, In *Proc. ISMIR 2004*.
- 7) Brian Whitman and Paris Smaragdis, Combining Musical and Cultural Features for Intelligent Style Detection, in *Proc. of ISMIR 2002*.
- 8) George Tzanetakis and Perry Cook, Musical Genre Classification of Audio Signals, *IEEE Transactions on Speech and Audio Processing*, 10(5), pp.293-302 (2002).
- 9) Jouni Paulus and Anssi Klapuri, Measuring the similarity of rhythmic patterns, in *Proc. of ISMIR 2002*.
- 10) Yusuke Tsuchihashi, Tetsuro Kitahara and Haruhiro Katayose, Using Bass-line Features for Content-based MIR", in *Proc. of ISMIR 2008*, pp.620-625.
- 11) Alex Sheh and Dan Ellis Chord Segmentation and Recognition Using EM-Trained Hidden Markov Models, in *Proc. of ISMIR 2003*.
- 12) Takuya Yoshioka, Tetsuro Kitahara, Kazunori Komatani, Tetsuya Ogata and Hiroshi G. Okuno, Automatic Chord Transcription With Concurrent Recognition of Chord Symbols and Boundaries, in *Proc. of ISMIR 2004*.
- 13) Amelie Anglade, Rafael Ramirez and Simon Dixon, Genre Classification Using Harmony Rules Induced from Automatic Chord Transcriptions, in *Proc. fo ISMIR 2009*.
- 14) http://www.music-ir.org/mirex/2005/index.php/Main_Page