

自動運転車のための自動音楽生成に向けて Towards a "realtime music generation" for autonomous vehicle

長嶋 洋一 *1
Yoichi Nagashima

*1 静岡文化芸術大学
Shizuoka University of Art and Culture

1. はじめに

世界中で開発が進められている自動運転車は安全走行のため多種の搭載センサを外界に向けており、GPSを含むこのセンシング情報を生かして周囲の状況に反応する「リアルタイム作曲システム」として、運転の必要がない全ての搭乗者に快適で音楽著作権の不要なBGMをライブ生成する自動作曲手法が求められている。英国Forkswagenが2013年に発表した“Play the Road” [1]は運転手目線のシステムでありやや目的が異なるが、その音楽スタイルの検討は重要な参考資料となった。ライブ生成の要素を除外としても、古今東西の音楽データを収集し深層学習させたAIが良質な音楽(BGM)を自動生成するためには、教師データとして膨大な音楽嗜好感性ビッグデータを適用すればよいが、いまだに成功例が報告されていない。

筆者は2006年にIPA未踏に採択され発表した“FMC3 (Free Music Clip for Creative Common)”において音楽的ヒューリスティクスを重視した考え方を提案した [2]が、豊田中央研究所に共同研究を依頼され開発した2016年版の車載用リアルタイムBGM生成システムにおいてこの原則を発展させた [3]。このリアルタイムBGM生成システムの試作に関して、(1)BGM進行のための音楽的な基本原理、(2)ループの繰り返しとリズム/ビートのスタイル、(3)センサ/マッピング/音楽生成のブロック分割、(4)音楽要素パラメータへの確率統計的な重み付け、(5)調性とコードとスケールの構成、(6)試作と実験の模様、などについては情報処理学会音楽情報科学研究会にて詳細に報告したので [4]、本稿では自動運転車が生成するセンサ情報と音楽音響的情報との関係について考察する。

2. 自動運転車のセンサ情報と自動作曲

最近の自動車では一般的な運転席からの視界のカメラ情報・CANデータ(運転状況センサ)・GPS情報などに加えて、自動運転車には、衝突防止/近接物体検出などの目的で多種のレーダセンサが搭載されており、漠然としたカメラ画像による画像認識と異なり、クルマに直接的に近接する物体や状況、すなわち運転車や同乗者が車窓の変化として感じる刹那的な状況変化を得ることが出来る。連続的に自動作曲演奏を続けるBGMとともに、このレーダセンシング情報に対応した刹那的なサウンド生成を行って、BGMシステムのサウンドとともに運転者/同乗者にトータルのサウンド(音楽)を提供することが望まれる。後者の刹那的に生成されるサウンドには、「楽音」と「効果音」という大きく2種類のタイプがある。

楽音とは「音楽的枠組みに乗っているサウンド」の事で、具体的には個々の音 (note) がピッチ (音高/音階) を持つもので、背景で鳴っているBGMの持っている調的枠組みに対して矛盾しないnoteを選ばないと、不快な不協和音になってしまう。例えばある瞬間にCM7という和音が鳴っていて、メロディー音もCM7に対して音楽的に選ばれたchord notesとtension notesが鳴っている時に、例えばF音とかF#音のように音楽的にavoid (禁止) とされる音が一瞬でも鳴

ることは避けなければならない。効果音とは「楽音以外のサウンド」であり、特定の音楽的ピッチ (音階) を持たないパーカッション (打楽器) 系の音、風や水やジェットなどのノイズ系の音、物体が衝突する打撃音や擦れる摩擦音、鐘や鈴の音や時計のカチカチ音、動物の鳴き声、人の話し声や叫び声や笑い声、ピッチを感じない重低音の衝突ノイズ、など多様な「騒音/自然音」が該当する。これらは基本的にBGMパートの持っている音楽的/調的枠組みとは異種なので、同時に鳴らされても人間は音楽的に不快な不協和音とは知覚しない。

3. リアルタイム・レーダセンシング・サウンド生成システムの構成

各種のレーダセンサに対応して「楽音」に属するサウンド (ジングルのようなある種の短いフレーズ/メロディー) を刹那的に生成するためには、上述のように背景で鳴っているBGMの持っている調的枠組みを考慮する必要があるが、BGMには調的枠組み (ピッチ方向) とともに時間的構造もあり、本システムでは基本的に4パターン (2小節×4回) という周期でループしている事を考慮する必要がある。つまり、レーダセンサから周囲の刹那的な変化情報が届いたその瞬間、BGMがいまだどこ (何小節目/何拍目) を演奏しているのか、という点を考慮する必要がある。ランダムに選ばれる4パターンのコード進行の途中であれば、「今鳴っているコード」「次に鳴るコード」という和声的情報を考慮すればいいが、4パターンのほぼ最後あたりの場合には、次の4パターンの最初のコードがまだランダム選択されていない瞬間には、その決定を待つまでイベント検出状態を保持する必要があり、イベント検出の直後にすぐサウンド生成を始めてしまわないようにする必要がある。

そこで本研究で試作開発したBGM自動生成システムに加えて、「リアルタイム・レーダセンシング・サウンド生成システム」には、「リアルタイム・レーダセンシング」・「パラメータ・マッピング」・「リアルタイム効果音生成」け・「リアルタイム楽音生成」が有機的に結合したシステム [5] を構築する必要があると提案した。

4. 謝辞

本研究にあたって、共同研究として実験/試作/検討を進めた、藤枝延維/赤井良行/宇田尚典 (豊田中央研究所ヒューマンサイエンス研究領域/感覚拡張プログラム) の各氏に謝意を表す。

5. 参考文献

- [1] <https://www.youtube.com/watch?v=3flwZ8OpXBY>
- [2] <http://nagasm.org/FMC3/>
- [3] <http://nagasm.org/ASL/paper/JASMIM2017.pdf>
- [4] 長嶋洋一, 自動運転車のためのリアルタイム作曲システムに向けて, 情報処理学会研究報告 (2018-MUS-118), 情報処理学会, 2018
- [5] <http://nagasm.org/ASL/paper/figure.png>