

統計言語モデルを利用した自動作曲システムのプロトタイプの構築

小林 正義[†] 山本 博史[‡]近畿大学大学院総合理工学研究科[†] 近畿大学理工学部[‡]

1. はじめに

一般的に作曲は難しそうというイメージが強く、実際に音楽知識の習得が必要となることで手の出しにくいものとなっている。しかしながら、音楽知識の習得に時間をかけなくとも、簡単な操作で作曲を行ってみたいと考えたことのある人は多いだろう。そこで、音楽知識の無いユーザを対象とした自動作曲システムの構築を最終目標として、本研究ではそのプロトタイプを構築する。具体的には、作曲方法としてコード進行を基とした作曲方法を採用し、統計言語モデルを利用して既存の曲の音楽的特徴を反映させることを試みる。そして、統計言語モデルにより自動生成した音楽的な文字列を実際に音に置き換えた場合に曲が成り立つのかを確認する。また、自動生成したコード進行を曲の一部として見なし、ユーザが複数の曲から自由に曲を組み合わせることで作曲を行うことのできる自動作曲システムのプロトタイプを構築する。

2. コードについて

コードとは高さの違う 2 つ以上の音を同時に鳴らした時の音のことであり、C(ドミソ)のようにコードは記号化される。このコードを連結した一連のコードの流れをコード進行と呼ぶ。

コードのうち、ダイアトニックスケール上に構成されるコードをダイアトニックコードと呼ぶ。ダイアトニックコードは音階を指定する調により異なり、ある調の間は基本的にその調のダイアトニックコードが使われる。ダイアトニックコードは響きの安定さからトニック(安定)、ドミナント(不安定)、サブドミナント(やや不安定)の 3 つの機能に分類できる。ドミナントからトニックに進行するように、響きが安定するようにコードを進行させることを解決と呼ぶ。解決された状態から再び解決する進行であるケーデンスを繰り返すことによるコード進行は作成される。

3. コードのモデル化

既存の曲のコード進行がケーデンスを基に作成されたものであると考え、既存の曲のケーデンスを反映させてコード進行を自動生成することを目的として統計言語モデルを利用したコード進行のモデル化を行う。

統計言語モデルとは、単語や言語表現の出現確率を求めるために利用される確率モデルのことである。コードは記号化され、コードを単語同様に記号として扱うことによりコードの並びの出現確率を求めることが出来るため、コード進行の生成が可能となる。

本研究では代表的な統計言語モデルである N-gram モデルを利用する。N-gram モデルは単語の生起確率が直前 N-1 つ前までの単語に依存すると仮定する統計言語モデルであり、文推定システムで推定した文が適切であるかを見ることに利用される他、文の生成システムにも利用される。

4. コード進行の自動生成

調によりダイアトニックコードが異なるため、学習データはある 1 つの調を用いる。また、音楽のジャンルでコード進行に違いが出ると考え、音楽ジャンルで学習データを分ける。

想定しているユーザは音楽知識の無い人であるため、特徴的な曲とするよりも自然で聴きやすい曲としたい。そこで、学習データのダイアトニックコードの確率のみを反映させ、ダイアトニックコードのみの構成となるコード進行を生成する。ダイアトニックコードはその調の音階から構成されるため、聴きやすさを重視できると言える。

コードの機能は 2 コード間で成り立つものであり、後続するコードの機能が直前のコードの機能に依存すると捉えることが出来るため、N-gram は次数 2 のものを用いる。一度も出現しなかった組み合わせについては 1-gram によるバックオフスミングを行うことにより補完する。本研究では、コードの生成時に乱数を求め、あるコードの次にくるコードの生起確率の和を見て、乱数が当てはまる確率範囲に該当するコードを次のコードとして決定、生成する方法を取る。これを指定の条件に当てはまるまで繰り返すことによりコード進行が自動生成できる。

5. 自動作曲システムのプロトタイプの構築

自動生成したコード進行を基に生成された曲を曲の一部と見なし、生成された複数の曲をユーザが選択、組み合わせることにより作曲を可能とする自動作曲システムのプロトタイプの構築を行う。

作曲の手順としては以下の通りとなる。

1. 曲の長さ、ジャンルに関する選択をし、作曲ボタンを押すと、選択に基づき 1 曲生成される
2. 生成された曲は複数曲保存でき、それぞれ曲の速さを選択でき、聴くことができる
3. 組み合わせたい順に保存された曲を選択すると、生成された曲を曲の一部とした新たな曲が作成できる

Building a prototype of automatic composition system using a statistical language model

[†]Masayoshi Kobayashi, Graduate School of Science and Engineering, Kinki University

[‡]Hirofumi Yamamoto, Department of Science and Engineering, Kinki University

6. 学習データの再現に関する実験

6.1. 実験内容

多くのデータが集められた音楽ジャンル・調であるハ長調のクラシック、マザーグース(英語圏の童謡)、琉球民謡を学習データとして使用する。

学習データのコード進行に対する生成されたコード進行の再現性を確かめるため、ダイアトニックコードのみで構成された一般的に使われるコード進行の出現しやすさを学習データから作成したモデルと生成データから作成したモデルを見ることで行う。評価には平均分岐数である perplexity(以下 ppl)を使用する。この値が大きいくほど後続する単語の種類も増えるため、単語列の予測、生成が難しくなる。今回は学習データを再現したいので、ppl 値が学習データに近い生成データのモデルを良いモデルと見なす。また、ppl は統計量に基づく値であるので、データ数を増やすことにより収束するという考えから、生成データのデータ数を増やすことにより信頼できる ppl 値が得られるものとして、生成データには学習データの 100 倍の曲数分コード進行を生成したものをを用いる。

ユーザが曲の長さを選択できるようにするため、生成されるコード数は任意の数にしたい。そこで、生成データの生成方法として 1 曲分のコード数を学習データの確率に基づき可変させる場合と、ある数に固定する場合とを比較し、コード数を固定する方法を用いることは適切であるかを検証する。コード数を可変させる場合は、学習データの 2-gram 確率に基づきコードの生成を終了させる方法をとる。あるコードの次にコードが後続しない場合の確率を構成コードに含め、この範囲に乱数が当てはまればコード生成を終了することでコード数を学習データに基づき可変させる。コード数を固定させる場合は、数を指定し、その数までコードが生成された時にコードの生成を終了させる方法をとる。この方法の値として、学習データの曲で最小のコード数、最大のコード数、最小と最大の間の中間のコード数を用いる。

6.2. 実験結果と考察

可変させる方法と固定させる方法 3 種類の結果を見て、コード数によってコード進行の出やすさが異なることが分かった。また、コード数を全体的に学習データに基づきコード数を可変させる方法よりも数を固定する方法の方が学習データのコード並びを再現しやすいという結果を得た。よって、固定する方法を用いるのは適切であると言える。可変させる方法をとった場合に精度が悪かった理由としては、可変させる場合の終了条件が甘く、生成毎のコード数にばらつきが出るからであると考えられる。実際に学習データと可変生成データのコード数に関する分散を比較したところ、全ての音楽ジャンルの可変生成データで学習データの分散の 2,3 倍程度の分散となった。また、琉球民謡を学習データとした場合はコード数を可変させる方法と固定する方法がほぼ同等の精度となった。琉球民謡の学習データのコード数の分散を調べると、他の音楽ジャンルと比べて分散が大きいため可変させた場合の精度が良かったのだと考えられる。

7. 文字列の楽曲化に関する実験と考察

7.1. 実験内容

統計言語モデルにより生成した文字列であるコード進

行を基に曲を作成した時、実際に曲として成り立つことを確かめる。曲はコード進行に基づいて作成した伴奏にドラムパートとしてハンドクラップを加えたものとし、自動作曲システムのプロトタイプを用いて再生する。1 曲分のコード進行を生成する際の選択肢に曲の長さとして学習データのコードの最小数、最大数、最小と最大の間の中間数を、音楽ジャンルとしてクラシック、マザーグース、琉球民謡を、速さとして BPM60, 120, 150 を採用した。

7.2. 実験結果と考察

自動生成したコード進行を基に伴奏として楽曲化することは可能となったが、聞いてわかる程度に学習データの音楽ジャンルの特徴を表現することは出来なかった。その理由として、コードに基づく伴奏だけでは音楽ジャンルの特徴を出すのに不十分であったからであると考えられる。音色を音楽ジャンルに合ったものに変えることや、更に生成したコード進行からメロディを生成できるようにすることにより、今よりも音楽ジャンルの特徴を表現できる曲を生成できるであろう。

8. 結論と今後の課題

今回、音楽知識のないユーザを対象とした自動作曲システムの構築を最終目的に、コード進行の自動生成とそれに基づく自動作曲システムのプロトタイプの構築を行った。ppl を用いた実験を行った結果、2-gram を用いたコード進行の自動生成を行う際に学習データの確率に基づきコード数を決めるよりも、コード数を任意に指定する方がより学習データを再現できることを確認した。また、生成したコード進行を基に楽曲化する際に、曲として成り立つが学習データの音楽ジャンルの特徴を表現するのに不十分であったため、音色の変化を付けることや、メロディを生成することにより特徴付けることが必要であることを確認した。また、今回は曲の再生は可能であるが、保存は出来ないため、音楽ファイル形式で作った曲を保存できる機能を付けることも今後の課題としたい。

参考文献

- 『決定版 コード進行スタイル・ブック』, リットーミュージック, 成瀬正樹(2011), p10-85
- 『同人音楽制作ガイド』, 秀和システム, 富井公・國田豊彦(2008), p90-122
- 『ケーデンス』, http://www.ab.auone-net.jp/~m-theory/contents/Music_Theory/cadence.html
- 『トニック、サブドミナント、ドミナントとは? / 初心者集まれ! 指板図くんのギター・コード講座 第 27 回 | 指板図くんのギター・コード講座 | ギター・マガジン』, <http://rittor-music.jp/guitar/column/guitarchord/463>
- 『続・トニック、サブドミナント、ドミナントとは? / 初心者集まれ! 指板図くんのギター・コード講座 第 28 回 | 指板図くんのギター・コード講座 | ギター・マガジン』, <http://rittor-music.jp/guitar/column/guitarchord/467>
- 『下ごしらえをしっかりと』, <http://vmei.sijen.biz/sakkyoku/sakkyoku04.html>