

歌詞からの音楽的要素の抽出

弘田修平[†], 吉田文彦[‡], 内田理[†], 菊池浩明[†], 中西祥八郎[†]

[†] 東海大学電子情報学部, [‡] 東海大学文学部

e-mail : sykes@kuwa.ep.u-tokai.ac.jp

1. はじめに

歌とは歌詞と曲を合わせることで音楽的な意味を持たせたものである。日本語文から自動的に作曲を行なわせる研究は、日本語文を小節に割り付け、そこからコード進行を自動生成するシステム[1]がある。このシステムでは入力された日本語文から音楽的要素の抽出を行なわない。また、どのような日本語文に対しても同様な出力結果しか得られないものである。歌詞の内容に関係なく、入力文の構成から曲の先頭と末尾以外の中間部分はコード進行を自動生成する。

歌詞から作曲する場合は歌詞の内容が重要であるが、筆者の知る限りでは歌詞に着目した自動作曲システムの研究はみあたらない。そこで、本稿では歌詞から曲を自動生成する為に必要な音楽的要素の幾つかを抽出する手法を提案する。

2. 音楽的要素抽出システム

2.1 システム概要

本システムの処理の流れを以下の図1に示す。

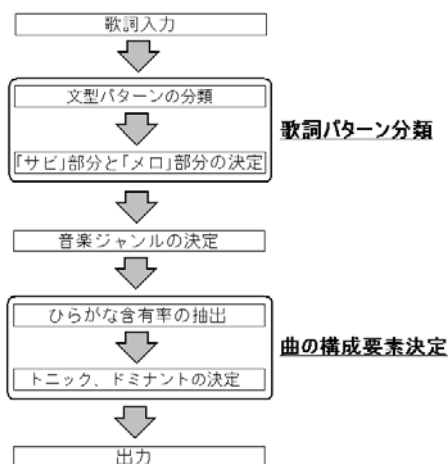


図1 処理の流れ

本システムは歌詞から作曲するための前処理として、入力された任意の歌詞から、作曲に必要な音楽的要素の幾つかを抽出するものである。

曲は「メロ」と「サビ」に対応する部分を持っている。そこで、それらを歌詞パターンとする。歌詞パターンとは歌詞中の文字の並びを図形と見立て、形と認識した時の様々な形の部分全てのことを指す(図2参照)。これらは曲を構成する上で最も重要な部分の1つである。そこで、歌詞パターンを複数持つという歌詞の特色により各パターンを判定し抽出する。次に、抽出した部分が「メロ」か「サビ」かを判定する必要がある

るのでこれらを決定する。

人間は「歌詞の雰囲気で作曲する」ということに着目し、雰囲気を顕著に表すブルース、演歌、こどもの歌、ヒップホップ、ポップスといった音楽ジャンルに分類する。

次に、歌詞の「やわらかさ」を抽出する。こどもの歌の多くは歌詞中にひらがなを多く含んでおり[5]、視覚上やさしい印象を受ける。また、やさしい印象を受ける曲は長調で構成されているという特徴を持っている。そこで、ひらがなの含有率により歌詞のやわらかさを抽出する。

最後に、歌詞中からコード進行を決定づける要素を抽出する。

2.2 歌詞パターンの決定

詩には自由詩と歌謡詞の二種類[5]がある。本システムでは作曲されることを前提とされている歌謡詞に対して処理を行なうものとする。

曲には図2に示すように複数に分かれた部分を持つ歌詞パターンという部分がある。

サビ部分とは、歌詞中で最も主張したい部分のことをいい、メロ部分はサビ部分以外のことをいう。歌謡詞はメロ部分とサビ部分に分けて書かれている。これらは字脚(ひらがなの文字数)が揃えられているので類似した文型パターンが多く見られる。文型パターンとは、曲を構成する歌詞パターン中の最小のパターン部分を指す。そこで、歌詞を文型パターンに分類する方法について次に述べる。また、分類された文型パターンの「メロ」部分と「サビ」部分の決定については2.3節で述べる。

2.2.1 文型パターン分類

異なる文型パターンが歌詞中に複数回出現するという歌謡詞の特徴から歌詞を図形に変換し、文型パターンに分類する。以下に文型パターンへ分類する方法を述べる。

- 1) 空行を検索し、空行があれば2)へ進む。空行がなければ3)へ進む。
- 2) 空行までを区切りとし、文型パターンと決定する(文型パターンには空行を含めない)
- 3) 歌詞中に漢字があればひらがなに変換する。促音は1文字とカウントし拗音はカウント数に含めない。スペース文字や記号も1文字とカウントする。英語の場合は1単語を1文字とし、記号へ置換する。
- 4) 変数 i 、 j 、 n 、 m 、 a を設定する。 i と j の初期値は1とする(図2参照)。

i = パターン候補先頭行

- j = パターン候補末尾行
 $m = j + 1 \dots$ パターン比較先頭行
 $n = m + (j - i) \dots$ パターン比較末尾行
 a = 全体行
- 比較ルーチンへ進み、結果が **false** なら次へ、そうでなければ 7)へ進む。
 - m と n を 1 増加し、5)へ戻る。
 - $j = (a/2)$ ならば次へ進む。 $j < (a/2)$ ならば j を 1 増加し 5)へ戻る。
 - A を文型パターンと決定する。決定した文型パターンとそれに類似したパターン部分を取り除く。このとき、各文型パターンの開始行と終了行の値を記憶しておく。続いて、歌詞の最初にあった文型パターンの末尾行の値を i と j へ代入する。また、 i から歌詞の末尾行までの行数の値を a に代入する。
 - 全てのパターンに分割されるまで 5)~8)を繰り返す。5)の処理を行う際に「 m = 記憶したいずれかのパターン(p とする)の開始行の値」であれば m に「 p の終了行+1」を代入し、再度処理を行う。6)の処理を行った際に「 n = 記憶したいずれかの開始行の値」になれば 7)へ進む。7)の処理を行った際に「 j = 記憶したいずれかの開始行」になれば 8)へ進む。
 - ひらがなに変換した歌詞を元に戻し、分割されたパターンとパターンの区切りに空行を挿入して処理を終了する。

比較ルーチン

$i \sim j$ 行 (以降、A とする) と $n \sim m$ 行 (以降、B とする) を比較する。A と B が類似していれば A をパターン候補とし、結果を返す。このとき、前のパターン候補があれば結果を置換する。類似していなければ **false** と結果を返す。類似は A の各行の文字数と B の各行の文字数が ± 2 文字までの差とする。

図 2 に文型パターンへ分類した例を示す(注 1)。

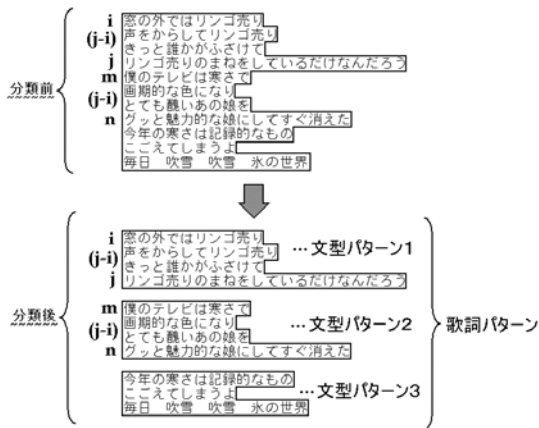


図 2 パターン分類の実行例

(注 1) 紙面の都合上、歌詞の一部と途中処理は省略した。

2.2.2 「サビ」部分と「メロ」部分の決定

(A) 「サビ」部分の決定

歌詞には必ずタイトルが与えられている。既存の歌詞を調査した結果、サビ部分中にタイトル文字列が含まれる場合が 83% (1000 曲中) [2]であった。そこで、「サビ」部分を決定する際はタイトル文字列に着目する。

(a) 文型パターンにタイトル文字列がある場合

文型パターン部分にタイトル文字列が含まれる場合はその文型パターンを「サビ」部分と決定する。

このとき、タイトル文字列と同じ単語を持った複数の異なる部分が存在する場合がある。この場合はサビ部分の判定が困難になる。以下に 2 つ以上のパターンにタイトル文字列が含まれる場合のサビ部分を決定する手順を述べる。

- タイトル文字列を検索語とし、すべての文型パターン部分を検索する。
- タイトル文字列があった場合は、6)へ進む。なかった場合は 3)へ進む。
- タイトル文字列を形態素に分け、タイトル文字列に構文解析[8]を施し、「受け」の最後の部分をサビ決定要素とする。

$$T = (m_1, m_2 \dots m_n)$$

$$\left(\begin{array}{l} T: \text{タイトル文字列} \\ m: \text{形態素} \\ n: \text{形態素数} \end{array} \right)$$

- m_1 を検索語とし、歌詞中から一致する文字列を検索する。一致する文字列が見つからない場合、 m_1 が名詞ならばその類義語を検索語とし、再度歌詞を検索する。類義語は同義語、狭義語、広義語の順で検索する。一致する文字列がなければ次の形態素を検索語とする。このとき、付属語なら次の形態素を検索語とする。検索語が検索途中で検出された場合は 5)へ進む。そうでなければ検索語が途中で検出されるか、 m_n になるまで繰り返す。
- 一致した文字列の行を構文解析し、「受け」の最後の部分とサビ決定要素を比較し、一致すれば 6)へ進む。一致しなければ 4)へ戻る。いずれも一致しなければ処理を終了する。s
- タイトル文字列がある文型パターンと、その文型パターンに類似したパターン部分をサビ部分と決定する。

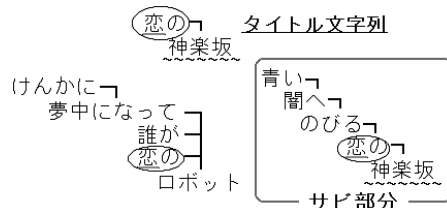


図 3 サビの決定例

図3は以上の手順によりサビ部分を決定した例である。

(b) 文型パターン中にタイトル文字列がない場合

歌詞中の最後に出てきた文型パターンをサビとし、その文型パターンに類似したパターン部分もサビと決定する。また、最後に出てきた文型パターンがどの文型パターンにも類似しない場合もサビ部分と決定する。

(B) 「メロ」部分の決定

サビ部分以外の文型パターンを「メロ」部分とする。メロ部分には複数存在することがあるので、異なる各メロ部分にはそれぞれ「A」「B」などの名前を割り当てておく。また、「A」と類似した「B」があった場合は「A」とし、作曲の要素に用いる。

2.3 音楽ジャンルの決定

(a) ブルース、演歌の決定

「ブルース」や「演歌」の決定には音数律の一つである七五調を用いて決定する。

七五調とは日本の詩歌・韻文などにおける音数律の一つである。俳句などの場合は七音節の句と五音節の句が続いたものを一単位としてこれを反復するものを七五調として扱うが、本システムでは五音節の句・七音節の句それぞれを一単位として扱うこととする。演歌とブルースを同カテゴリーとした理由として、音楽理論の観点から見て共通点が多くみられるという点が挙げられる。

五音節および七音節の決定手順は以下に示す。

- 1) 形態素へ分解し、漢字をひらがなへ変換する。
- 2) 各行中にスペース文字があればそこまでを音数律とする。このとき、文頭からスペース文字までが9文字以上なら3)へ進む。
- 3) 行頭から形態素の組合せを五音節または七音節になるよう区切る。このとき、区切れの先頭は自立語になるようにし、付属語の前で区切らないようにする。(内容展開度合いの抽出[1]に基づく) 組合せを調整する際、a、b、cを形態素とした{a, b, c}の文章があったとする。そこから{a+b, c}または{a, b+c}のように、どちらの組合せでも七五調になり得る場合がある。このとき、{a, b, c}の中心であるbが付属語ならば a+b で七五調を決定し、自立語なら b+c で七五調を決定する。また、cが付属語であれば b+c を七五調と決定する。

図4に音数律の組み合わせ例を示す。

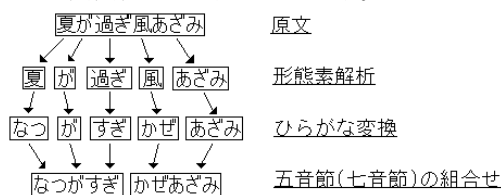


図4 音数律の組み合わせ

続いて五音節または七音節の数をカウントし、歌詞全体の音数律の中から七五調の割合を算出する。七五

調の割合が閾値を超えればこの歌詞を「ブルース」「演歌」と仮決定する。次に、1つのメロ部分と1つのサビ部分を持つ歌詞パターンを一単位として三回以上出現する歌詞は「演歌」とし、それ以外を「ブルース」とする。また、演歌に英語が含まれることはない[2]ので英語が歌詞中に存在しない場合は「演歌」とし、英語がある場合は「ブルース」とする。既存の曲の解析結果より、ブルース・演歌に含まれる七五調の割合は70%以上[2]であった。このことから、演歌・ブルースを判定する閾値を70%と決定した。

(b) こどもの歌の決定

「こどもの歌」を決定するには歌詞中の七五調の割合とひらがなとカタカナの含有割合を用いて判断する。便宜上「こどもの歌」としているが、音楽的には「明るい印象を受ける曲全て」のことを指すこととする。七五調の割合算出方法は「ブルース」や「演歌」と同様とする。既存の曲の解析結果より、こどもの歌に含まれる七五調の割合は70%以上、ひらがなとカタカナを合わせた含有割合が65%以上[2]であった。このことから七五調の割合が70%以上であり、同時に歌詞中のひらがなとカタカナを合わせた含有率が65%以上の歌詞を「こどもの歌」と決定する。

(c) ヒップホップの決定

ヒップホップはメロ部分で「韻」を踏むという特徴をもっている。また、類似するメロ部分を持たない(例えばA、B、Cと類似したメロが一つもないもの)場合が多い。そこで、類似したメロ部分を持たずに、各メロパターンが「韻」を踏んでいればヒップホップと決定する。韻を踏んでいるかどうかを判定するには以下の手順で判断する。

- 1) メロ部分をひらがなに変換する。
- 2) 1行中にスペース文字があればそこまでを一区切りとし、各区切りの末尾2文字の母音を比較する。各末尾の母音が一致すれば「韻」と判断する。スペース文字がなければ3)へ進む。
- 3) i行目と2行目の末尾2文字の母音を比較し、一致すれば「韻」と判断する。

(d) ポップスの決定

ここまでで音楽ジャンルを決定できないものはポップスと決定する。理由は、ポップスはロック・R&B・フォーク・ダンス音楽など様々なジャンルの音楽性を内包していることが挙げられる。また、その内容は時代によって変動する。そこで、それらに幅広く対応できるジャンルとしてポップスを定義し、そこへカテゴリー分類する。

2.4 歌詞中のひらがな含有率の抽出

日本語には漢字・カタカナ・ひらがなが混合(場合により英語も含む)して文に存在するという特色を持っている。平安時代に女性が自己表現の方法としてひらがなを用いた。そこで、女性によるやわらかな表現はひらがなによって成り立つと考え、ひらがなを多く含む歌詞を「やさしい」とする。以上よりひらがなに

着目し、その含有率によってやわらかさを決定する。

既存の歌詞 1000 曲[2]から 1 曲あたりの平均混合比率を調査した結果、ひらがな 55%、漢字 22%、カタカナ 5%、英数字 18% (うち 17%は英語) で構成されていることがわかった。また、55%~70%までの間の含有率をもった歌詞は少なかった。このことより、ひらがなの含有率の閾値を 65%とする。閾値よりひらがなの含有率が高ければ「やわらかい」歌詞と判断し、明るいイメージである要素が強いとする。

2.5 トニック・ドミナントの決定

聴覚上落ち着きのあるコードであるトニック (主和音) と不協和音であるドミナント (属和音) が決定すればコード進行[4]が決定される。そこで、Positive な語をトニック、Negative な語をドミナントと割り当てる。また、ドミナントの多くは誤って決定されるので、不安定に感じさせる和音のサブドミナントも Negative 語と割り当てる。

割り当てた Positive な語と Negative な語それぞれは TeX-Ray(注 2)を用いて抽出する。TeX-Ray は形態素解析、構文解析、概念の抽出、PND (Positive - Negative Dictionary) による Positive 語、Negative 語の判定と抽出などを行うことができる日本語解析ツールである。本システムでは、TeX-Ray の PND による Positive 語、Negative 語の抽出機能を用いた。否定の文である場合は正しい判定ができないので、否定文がある場合は抽出された単語を Neutral とする。判断するには抽出された行の構文解析を行ない、単語後に否定する語があれば否定文であると判定し Neutral (トニック、ドミナント、サブドミナントのいずれか) と決定する。また、Positive と Negative 以外の語は Neutral とする。

3. 評価と考察

3.1 評価

実際に歌詞から音楽的要素が抽出できるか判断するために実験を行った。実験に用いた歌詞は、異なるジャンルの作曲された既存の曲 10 曲[6]を選出した。以下に本システムで抽出した音楽的要素の出力結果と実際の曲を比較した結果を以下に示す。

- ・文型パターン決定の正解率=90%
- ・サビ部分・メロ部分決定の正解率=70%
- ・音楽ジャンル決定の正解率=80%
- ・やわらかさ抽出の正解率=70%
- ・トニック・ドミナントの正解率=56%

3.2 考察

文型パターン決定は 90%と、ほぼ全てのパターンを本システムで分類することができた。それに対し、サビ部分・メロ部分の決定は 70%と、20%が正しく決定できない結果となった。理由として、タイトル文字が英語に対し歌詞中には和訳された単語が入っていたこ

(注 2) 日本語解析器「TeX-Ray」, 東海大学文学部広報メディア学科, 吉田文彦作

とが挙げられる。そこで、タイトル文字列の単語を和訳しサビ部分の特定を行ったところ正解率は 80%になった。残りの 10%は人間でも判断が難しいものであった。これについては連想される語すらサビ部分に入っておらず、歌詞を読んだ上で人間が思考しなければサビ部分を判定することができないものであった。こういった歌詞については、別のサビ決定方法を考案する必要があると考えられる。

音楽ジャンルの決定は殆どが正しく行われた。また、ジャンル決定に文字の含有率を用いることは有効な方法だと言える結果となった。完全に分類できなかった理由として、定義が幅広いポップスに分類されてしまったことが挙げられる。ポップスに分類されたものの多くは人間であればより細かく分類することができる。これを分類するには各ジャンルの歌詞の特徴を調査する必要があると考えられる。それにはまず、未調査である英語 (数字)、カタカナ、漢字それぞれの含有率に着目すれば幾つかのジャンルに分類することが出来ると考えられる。

トニック・ドミナントの正解率はまずまずの結果となった。決定したコードの半数が誤った理由は、歌詞に関係なく音楽理論によってコードが決定されているからである。しかし、正解した語の多くは曲の構成と密接な繋がりが見受けられた。また、歌詞に 100%依存することなくコードが決定されているということも結果から判断することができる。つまり、作曲する上で制限が少ないと言える。このことから、作曲する際に必要な曲を構成する要素を抽出、決定することができ、自由度の高い作曲ができると考えられる。

4. まとめ

歌詞から音楽的要素の幾つかを抽出する方法を述べた。今後は、ポップスに含まれる音楽ジャンルを細分化し、そこから作曲する際に必要な音楽的要素を抽出する予定である。また、抽出した音楽的要素より実際に自動作曲するシステムを構築することも今後の重要な課題である。

参考文献

- [1] 早川, 稲垣, 田中: 4G-4 歌詞がメロディと和音に与える影響の検討, 情報処理学会全国大会講演論文集, Vol. 第 59 回平成 11 年後期 Num. 2 pp."2-7"- "2-8" 1999.
- [2] うたまっぷ: <http://www.utamap.com/>
- [3] 長尾真: 岩波講座ソフトウェア科学 自然言語処理, 岩波書店, 1996.
- [4] 北川祐: ポピュラー音楽理論, リットーミュージック, 1994.
- [5] 阿久悠: 作詞入門, 産報, 1972.
- [6] 株式会社ライトスタッフ: フォーク・ロック 1001, ドレミ楽譜出版社, 2000.
- [8] 奈良先端科学技術大学院大学情報科学研究科松本研究室: 日本語係り受け解析システム「南瓜」, <http://chasen.org/~taku/software/cabocha/>