

歌謡曲における「さび」の楽譜情報に基づく特徴抽出 — 小室哲哉の場合 —

村松 純

NTT コミュニケーション科学基礎研究所
pure@cslab.kecl.ntt.jp

小室哲哉氏作曲の楽譜の旋律情報から、「さび」の特徴抽出を試み、抽出した特徴に基づいた「さび」認識システムの性能を評価した結果を報告する。

Extraction of Features of 'Sabi' from the Score in Japanese Popular Music Songs — in the Case of Tetsuya Komuro —

Jun Muramatsu

NTT Communication Science Laboratories
pure@cslab.kecl.ntt.jp

The extraction of features of 'sabi' from the information about the melody in the scores, which are composed by Tetsuya Komuro, will be studied. The performance of 'sabi' decision system based on the extracted features will be reported.

1 はじめに

美しく魅力的な曲を作るためには旋律にどのような特徴があれば良いかを解明するために、単旋律で音域が限定されている歌謡曲の「さび」に注目して研究を進めている。歌謡曲をミリオンヒットをさせるためには印象深い「さび」を持つことが必要条件であると言われている。多くの人に受け入れられる良い「さび」を作るためには、作曲の技術以外の感性が必要であると思われる。

本研究では、「さび」とはどのような特徴を有しているのかを探求するために小室哲哉氏作曲の楽曲における「さび」の特徴抽出を試みる。本来ならば、伴奏や歌手が歌う

ときの力の入り方等の細かな演奏情報も「さび」の認識に大きくかかわっている可能性があるが、ここでは、楽譜(ボーカスコア)から得られる旋律の情報のみから特徴を抽出する。そして抽出した特徴を用いて計算機に「さび」を認識させることによって特徴抽出の妥当性を検証する。

平田、青柳は[1]において、黑板モデルを用いて「さび」の認識システムを構築する提案をしている。そこでは「転調がある」「さびでは高い音が使われている」等の経験的に良く知られている特徴があげられている。平田、青柳の研究では認識ルールを人間が与えるアプローチであるのに対して、本研究では「さび」の認識ルールを事例から学習させるアプ

ローチである点が異なっている。

本研究では4/4拍子の楽曲のみを扱う。そして以下では全てのフレーズをハ長調に移調して議論し、転調やコード進行には直接的には立ち入らない。また、平田、青柳は[1]において、楽曲をフレーズに分割するところから議論をはじめているが、本研究ではフレーズの区切りはあらかじめ与えられていることを前提として話をすすめる。今回は楽曲をフレーズに分割する作業は人間の手作業で行っている。

2 人の「さび」の認識に関する調査

「さび」というのは本来楽曲の構成から決定されるもの([2]参照)であるが、その部分を盛り上げたり印象深くなるように作曲されることが多いため、音楽の専門知識がない人でも同じ箇所を「さび」と判断すると予想できる。そこで、「さび」の認識に個人差が現われるかどうかの調査を行った。

21人の被験者(20-25歳)に小室哲哉氏作曲の楽曲(12曲)を聴かせ、「さび」の箇所と、その理由を尋ねてみた。このとき「さび」がどんなものであるかを明示せず、個人の主観で判断してもらった。その結果、20曲で9割以上の方が同じ箇所を「さび」であると答え、その曲を知らない人(非常にわずかな数であったが)も知っている人と同一の所を「さび」であると答えた。また、「さび」である部分を知っていると答えた場合を除けば、9割以上の方が全ての曲に対して「さび」の理由を「盛り上がっている部分(クライマックス)だ」もしくは「印象的な部分だ」と答えた。

しかしながら、「メロディ」「リズム」「歌詞」「演奏(歌い方や伴奏)」の各要素に対して「さび」には他の部分にはない特徴が

あるかどうかを尋ねたところ、答えに個人間のばらつきが見られた。また、一人に対して各曲の答えにもばらつきが見られた。

以上のことから、「さび」は「印象深い部分」「盛り上がっている部分」と認識されているが、どこに特徴があるかの認識には個人差や曲毎の差が大きいことが分かった。

3 旋律の特徴

旋律は音を出すタイミングや長さ等のリズムに関する情報と、音の高さの情報からなる。本研究では一つの音符(休符)に対して、以下の四つの属性に注目した。

- Beat** 小節内の発音開始時間。すなわちその音が小節内の何拍目で始まるかに関する情報。小節の長さを48として0-48の数字で表す(例:拍が24であるとは、音符が2拍目に発音を開始されたことを意味する)。
- Duration** 音価、すなわち音の持続時間。4分音符の長さは12とする(例:音価が6であるとは、それが8分音符であることを意味する)。
- Interval** 音程。すなわち前の音からの高さの変化。ただし直前が休符だった場合はその前の音からの変化とする。半音でいくつ音が上下したかを整数で表し、休符の場合は‘r’と記す(例:音程が-2であるとは、直前の音に対して2半音(すなわち全音)下がったことを意味する)。
- Scale** 音階。ただしオクターブの違いを同一視する。ドの音を0として数字を順に割り当て0-11の数字で

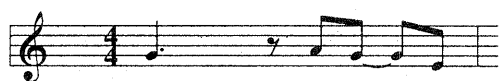
表し、休符の場合は 'r' と記す (例: 音階が 9 であるとは、ハ長調に移調したときにそれがラの音であることを意味する)。

旋律は音符の系列であるので、複数の音符の列を一つのまとまりとして考えることもできる。以下音符列について、上記の属性に関する条件を課したものを「旋律の特徴」と呼ぶ。ここで、1 個の音符も、個数 1 の音符列として扱う。

音符列を以下のように記述する。一つの音符は上記の四つの属性を持つので、これを 'x' で区切って順に並べる。図 1 の例において、2 拍目の 8 分音符は '42:6:-2:4' と記述される。

もしもある属性に注目しない場合はその属性の値の違う音符列を同一視して扱い、そこに何も記述しない。図 1 の例において 3 拍目の 8 分音符も 3.5 拍目の 8 分音符も ':6::' と記述され、それが何拍目にどのくらいの音程で、どの音階の音が発生したかということとは区別されない。

2 個以上の音符列からなる場合はそれを 'x' で区切って連結する。図 1 の例において、1 小節全体からなる音符列の音階のみに注目した場合は、':::7|:::r|:::9|:::7|:::4' と記述される。



拍: 1 2 3 4
 対応する数字: 0 12 24 36

図 1: 楽譜の例

4 「さび」の認識に関する二つの立場

旋律から「さび」を認識する時に次の二つの立場が考えられる。

立場 A 一つのフレーズを与え、それが「さび」であるかどうかを決定する(「さび」の絶対評価)。

立場 B 一つの楽曲を複数のフレーズに分割し、それらを比較して「さび」のフレーズを決定する(「さび」の相対評価)。

実際に人間が上記のどちらの立場に立って「さび」を判定しているかは今のところ定かではなく、おそらくこれらが組み合わされたものであると推測される。

5 立場 A に関する調査

この節では A の立場に立って議論を進める。

旋律は音符の時系列であることから音符の発生(遷移)確率を用いて楽曲分析がなされることが多い([3],[4]参照)。しかし、本研究では「音符がフレーズで使用される確率」を議論する。ある音符からある音符へ遷移するかどうかに関する情報は注目する音符列の個数を 2 個以上にすることによって知ることが出来る。

まず、音符が「さび」で用いられる確率と「さび」以外のところで用いられる確率について、39 曲の 125 フレーズ(そのうち「さび」は 39 フレーズ)を標本数とした有意水準 1% の二項分布の検定を行った。

表 1 は抽出された特徴の数を示した表である。第 1 列は注目した音符の属性の種類で、記号は前節で列挙した属性の頭文字を表している。そして第 1 行は音符列の長さを表している。数字は有意差が認められた特徴の数である。例えば表で BD の行の個数 3 の列の

数字 7 は、拍と音価に注目した連続する 3 個の音符列の中で、7 種類の音符列に有意差が見られたことを意味する。

一般に音符列の個数が少なければ特徴の種類が少なくなり、どのフレーズでも使用されるために有意差がでにくくなる。逆に、音符列の個数が多くなると、特定のフレーズでしか用いられなくなり、現時点の標本数では有意差が出る程の特徴にはならなくなると考えられる。また同じ音符列の個数のところで見た場合、リズム (拍と音価) のみに関する特徴よりも音の高低のみに関する特徴のところで多くの有意差が見られることから、「さび」はリズムよりもむしろ音の高低に特徴があると考えることが出来る。

属性	連続する音符列の個数							
	1	2	3	4	5	6	7	8
BD	0	3	7	9	9	4	1	0
D	1	2	5	8	9	9	8	4
B	0	0	1	5	8	8	4	2
IS	7	8	12	17	6	2	0	0
I	1	4	11	15	6	1	0	0
S	0	3	6	13	14	3	2	0
BDIS	11	12	2	0	0	0	0	0
BDI	3	11	1	0	0	0	0	0
BDS	5	15	10	1	0	0	0	0
DIS	4	13	6	0	0	0	0	0
DI	2	8	4	1	0	0	0	0
DS	3	14	10	2	0	0	0	0
BIS	11	13	4	1	0	0	0	0
BI	4	6	2	1	0	0	0	0
BS	4	12	10	2	0	0	0	0

表 1: 抽出された特徴の数

次にこれらの特徴のうちの一つのみに着目して「さび」判定を試みた。ただし、今回の調査では 125 フレーズ中に「さび」は 39 フ

レーズしか含まれていないため、常に「さびではない」と答えても 68.8 % の正解率がある。表 2 は全体の正解率の高かった特徴を列挙したものである。ここで、「+/-」の欄に「+」と書かれている場合は、その音符列が「さび」で多用されていることを表し、「-」と書かれている場合は、「さび」であまり使われていない音符列であることを表している。正解率の上位のものはどれも多用されている音符列であった。

どの特徴にも音程や音階の情報が含まれていることから、ここでも「さび」にはリズムよりも音の高低の特徴が大きく現われていることが分かる。また、判定誤りのパターンを調べたところ、「さび」でないものを「さび」であると判断する確率よりも「さび」であるものを「さび」でないと判断する確率が比較的大きかった。

特徴	+/-	正解率
0:12::0	+	75.2%
0::-4:0	+	75.2%
12::3:0	+	74.4%
18:6:2:4	+	74.4%
18:12::0	+	73.6%
:12:0:0	+	73.6%
0::-7:	+	73.6%
::r:r :-7:9	+	73.6%
:::9 :::0 :::2	+	73.6%
18::2:4	+	73.6%
18::2:4 24::0:4	+	73.6%
:6:2:2 :12:2:4	+	73.6%

表 2: 「さび」判定の正解率の高い特徴のリスト (上位のみ)

6 立場 B に関する調査

この節では B の立場に立って議論を進める。この立場に立った場合、転調等の情報も「さび」判定では有効であると思われるが、最初に述べたように、本研究では転調に関しては直接は立ち入らずに旋律の特徴に限定した調査を行う。

本節では、一つの楽曲をあらかじめ複数のフレーズに分割しておき、そのいずれかの部分を「さび」と判定させる問題を考え、その正解率を議論することによって特徴抽出の妥当性を検証する。

今回の調査では一つの楽曲中のフレーズの種類は 2-6 個の幅に入っており、でたために答えたとしても平均的に 33.2 % の確率で正解を出すことが出来る。

立場 A ではフレーズ単体で評価するため、「楽曲中で最高音をとる」等の特徴を利用することができなかったが、立場 B ではそれを利用することが出来る。そこで、「楽曲中の最も高い音」をとる部分を「さび」と判定する場合の正解率を調べたところ、正解率は 55.6 % となった。この事実は「高い音が出現するところがさびである」という「さび」について良く言われることがそれほどは当てはまらないことを意味する結果となっている。

表 3 は注目する音符の属性を限定して有意差のあった特徴を集め、一つの曲の中で最も多くの特徴を持ったフレーズを「さび」と断定したときの正解率の表である。

ここで正解率の算出には一つ抜き法 (cf. [5, pp.69-70]) を用いた。今回の場合は、まず 39 曲を 38 曲と 1 曲に分け、まず 38 曲で学習を行い、残りの 1 曲の「さび」判定を行ってテストする (ここで残りの 1 曲は特徴抽出に用いない未知の曲であることに注意)。これをテストの曲を変えながら 39 回行

い、平均の正解率を算出する。表にある正解数は 39 回のテストの曲に対して正解を一意に決定した数で、不正解数は完全に不正解になった曲数である。また正解率の計算では、「さび」の候補となるフレーズが N 個あった場合正答曲数を 1/N 曲と数えている。

これによると、拍、音程、音階の 3 つ (BIS) に注目して特徴を集めたときに最も正解率が高い (63.7%) ことが分かる。

属性	正解数	不正解数	正解率
BD	7	14	34.4%
B	14	17	45.3%
D	3	15	26.3%
IS	17	16	51.3%
I	13	15	45.1%
S	16	18	47.4%
DIS	14	12	48.7%
DI	2	17	23.7%
DS	17	9	57.5%
BDIS	14	15	46.4%
BDI	8	18	33.8%
BDS	18	17	50.4%
BIS	20	9	63.7%
BI	14	15	47.9%
BS	14	17	45.1%

表 3: 「さび」判定の正解率

7 まとめ

小室哲哉氏作曲の楽譜の旋律情報から、「さび」の特徴抽出を試み、「さび」判定システムの性能を評価した。今回の結果が一般の歌謡曲に付いて言えるかどうかは他の作曲家の楽曲も含めて調査する必要があると思われる。

本研究ではある特徴を持つ音符列が「使用される確率」を議論し、「遷移する確率」を議論しなかった。旋律からより細かい情報を得るには、適当な確率モデルを考えて遷移確率を議論する必要があると思われる。

本研究の方法は「さび」の特徴抽出の問題に限らず、例えば、作曲者の特徴抽出¹や「好きな曲」「嫌いな曲」等の人の好みの特徴抽出にも応用することが考えられる。このようにして抽出した特徴を多く有している旋律もしくは抽出した特徴を用いて生成した旋律が果して意図したものになるのかどうかということに関しては、今後の課題として残されている。

謝辞

本研究を行うにあたり、NTTコミュニケーション科学基礎研究所の向内隆文社員、平博順社員、藤本和則研究主任には研究のヒントと多くの助言を頂きました。白木善尚主任研究員ならびにメディア表現研究グループの皆様には多くの助言を頂きました。石井健一郎企画部長には本稿に目を通して頂き、貴重なコメントを頂きました。小山特別研究室の皆様には研究を進めるにあたり多くの助言、協力をして頂きました。この場を借りて感謝申し上げます。最後に、本研究を共同で

進める途中で急逝された故小山謙二特別研究室長の生前の御指導に謹んで感謝の意を表します。

参考文献

- [1] 平田圭二, 青柳達也: 普通の国のありふれた曲の認識システム — アリフゼロ —, 日本ソフトウェア科学会第4回大会論文集, pp.17-22, 1987.
- [2] 音楽之友社: 新音楽辞典 — 楽語 —, 音楽之友社, 1977.
- [3] 村尾忠廣: 楽曲分析における認知, 「音楽と認知」, 第1章, 東京大学出版会, 1987.
- [4] 村尾忠廣: 音楽の分析・解釈, 「コンピュータと音楽の世界」, pp.216-223, 共立出版, 1998.
- [5] 石井健一郎, 上田修功, 前田英作, 村瀬洋: 分かりやすいパターン認識, オーム社, 1998.
- [6] 村松純, 小山謙二: リズムパターンに注目した作曲家の特徴抽出, 日本感情心理学会第7回全国大会, 1999.

¹拍と音価からなるリズム情報のみから作曲者の特徴抽出をする試みとして村松, 小山 [6] の報告がある。