

暗意実現モデルの拡張に基づく楽曲間類似度の評価

Evaluating Melodic Similarity based on an Extended Implication-Realization Model

矢澤櫻子^{*1}
Sakurako YAZAWA^{*1}

浜中雅敏^{*2}
Masatoshi HAMANAKA²

宇津呂武仁^{*3}
Takehito UTSURO^{*1}

^{*1} 筑波大学大学院システム情報工学研究科知能機能システム専攻

Graduate School of Systems and Information Engineering University of Tsukuba #1

^{*2} 京都大学大学院医学研究科 ^{*3} 筑波大学システム情報系知能機能工学域

Department of Clinical System Onco-Informatics Kyoto University #2

Faculty of Engineering, Information and Systems University of Tsukuba #3

This paper proposes to measure similarity of melodies based on Implication-Realization Model (IRM), a music theory that abstracts music and then expresses music through symbol sequences based on information constituting the music such as pitch, rhythm, and rests and so on. This paper especially extends IRM so that the theory becomes much more appropriate to measuring similarity of melodies. More specifically, compared with the symbols of the original IRM, we introduce finer grained symbols by simply distinguishing up and down of interval directions and by dividing each most symbols of the original IRM into two extended symbols. Furthermore, we implement a parser which transforms tone sequence of an input melody into a sequence of the extended IRM symbols. The results of experimental evaluation through subjective human judgments show that the proposed extended IRM symbols outperform the original IRM symbols with respect to measuring similarity of melodies.

1. はじめに

本稿では音楽理論暗意実現モデル [Narmour 1990] [Narmour 1992] に基づいた楽曲間類似度計算について述べる。暗意実現モデルでは楽曲をシンボルと呼ばれる記号列に抽象化し表現することができる。今回は特に、楽曲間類似度計算に適した暗意実現モデルの拡張について述べる。

本稿では、音楽理論に基づいた分析結果を楽曲間類似度計算に用いる方法を検討した。楽曲間類似度計算の研究はさかに行われており、ユーザーに好みの楽曲を選ばせ、その楽曲情報からユーザーの嗜好を表すベクトルを生成し好みの楽曲を類似楽曲として選ぶ手法 [Hoashi 2003] や入力に音響特徴量を用いたニューラルネットを構築することで類似楽曲を抽出するシステム [Lampropoulos 2004]、離散フーリエ変換をかけてフレーズのパターンを見いだす手法 [Velardea 2013]、楽曲に含まれる音符列を文字列に変換し、その文字列に対して N-gram を適用することで類似楽曲の抽出を行う試み [Doraisamy 2004] など様々な手法が提案されている。

音楽理論を用いた従来研究としては、音楽理論 GTTM [Lerdahl 1983] に基づいた楽曲間類似度計算方法があげられる。GTTM はメロディに含まれる音の重要度を計算し、タイムスパンツリーと呼ばれる木構造を構築することでメロディを階層的に捉え分析、理解する音楽理論である [Hamanaka 2007]。

GTTM に基づいた楽曲間類似度計算 [Hirata 2013] は GTTM におけるタイムスパンツリーが楽曲間で合致している部分に対して、そのタイムスパンツリーに付いている音の長さ・数の位置にある音の総数を用いて類似度を計算していた。この手法はほぼ同じ曲の場合にのみ類似計算ができるものであり、同

じ曲を検索する場合にはとても適している。しかし制約がとて厳しいため、人間が聴いた場合に「ある程度似ている」と感じる曲の組の類似度が十分高くないという問題があった。

本研究では音楽理論暗意実現モデルを用いた楽曲間類似度計算法を提案する。暗意実現モデルを用いた先行研究として [Grachten 2005] は暗意実現モデルで定義されている構造を用いる手法を提案し、類似度を計算できることを示した。しかし、暗意実現モデルは本来、楽曲間類似度を計算するための音楽理論ではないため、オリジナルの構造を用いただけでは、類似度計算という観点では精度が下がってしまう問題があった。そこで我々はオリジナルの暗意実現モデルに対して拡張を行うことで楽曲間類似度計算精度の向上を図った。

我々はオリジナル暗意実現モデル、および、拡張を行った暗意実現モデルの両方で楽曲間類似度計算を行った。計算結果に基づき類似楽曲をデータベースから選び出し、被験者主観実験を行った。その結果、拡張を行った暗意実現モデルでの類似度計算がオリジナルの暗意実現モデルを上回った。

2. メロディへの暗意実現モデルのシンボル付与

2.1 暗意実現モデル

楽曲間類似度判定システムを構築するにあたり、暗意実現モデルという音楽理論に着目した。暗意実現モデルとは音楽学者である Eugene Narmour によって提唱された理論である。この理論音同士の関係をネットワークのように捉えてメロディを記述しており、このネットワークを解析・抽出することにより、メロディがどのような構造になっているかを分析することができる。また実際の解析は楽曲を構成する音高、音程、リズムや休符等の情報を用いて楽曲をシンボルと呼ばれる記号を用いて記号列へと抽象化して表現することができる。

連絡先: 矢澤櫻子, 筑波大学大学院システム情報工学研究科
博士後期課程, sakurako@music.iit.tsukuba.ac.jp

2.2 暗意実現モデルの拡張

暗意実現モデルを用いた楽曲間類似度判定は先行研究より、シンボルを用いた記号列比較によって類似度計算がある程度可能であることを示した。しかし、オリジナルの暗意実現モデルで定義されているシンボルでは音程の上下方向の仕方についての定義が無いために類似度計算という観点では精度が下がってしまっていた。例えば音程が「下がって上がる」と「あがって下がる」という振る舞いをしている音列が楽曲に出て来た場合に、全く違う旋律を我々は聴いているのに同じシンボルが割り当てられてしまっていた。

類似度計算に主眼をおいた場合、先のような聴いた印象が全く違う旋律音列が同じシンボルと判定されると、楽曲の展開は同じでも実際に聴いた場合に似ていないと判定されてしまうという問題が発生する。この問題を解決するため我々は音程の上下方向への定義の拡張を行った。拡張の具体的な例を図 1 に示す。

2.3 拡張暗意実現モデルのシンボルパーサー

本節では構築した暗意実現モデルパーサーについて述べる。本パーサーでは入力した midi 形式の楽曲を分析し、暗意実現モデルのシンボル列を出力する。

パーサーはまず、音価が変化する部分と休符を切れ目として音列のグループを作る。そのグループの先頭から連続する3音ずつをさらに細かいグループとし、連続する2つの細かいグループは1つの音を共有する。つまり、連続する2つのグループに含まれる連続する5音のうち前半3音と後半3音に対してシンボルをひとつずつ割り振る。この音列をシンボル化する際の音列の音数の場合分けを行うと以下の4通りとなる。1) $k=1$, 2) $k=2$, 3) $k=2n+1$ ($1 \leq n$), 4) $k=2n+2$ ($1 \leq n$), ここで k は割り振るシンボルを示している。1)と 2)の場合、含まれる 1 音もしくは 2 音に対して例外型と呼ばれるシンボルを割り振る。3)の場合は全ての組み合わせにシンボルを割り振り、4)の場合は最後の 2 音に対してのみ例外型を割り振る。

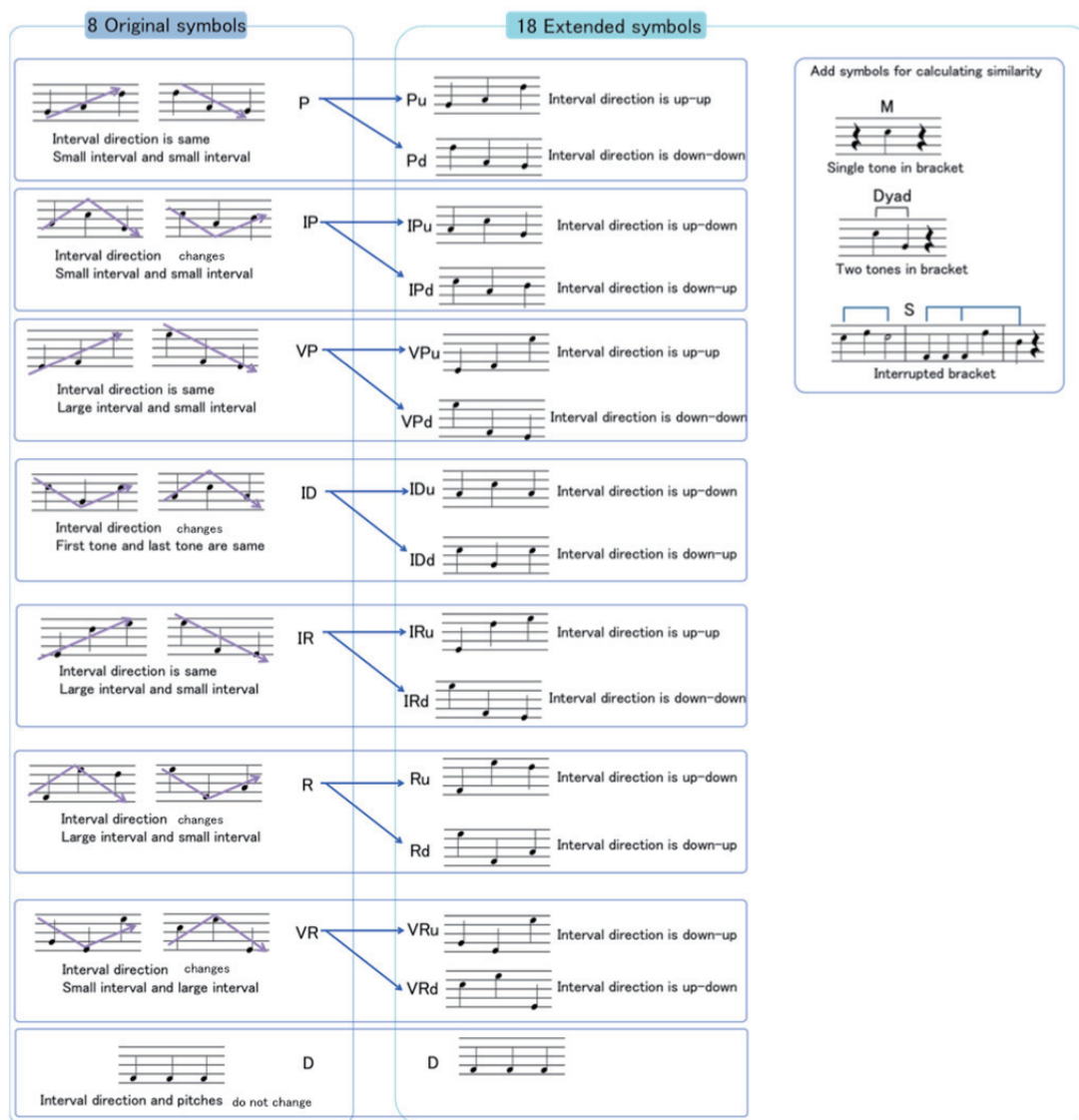


図 1, 拡張暗意実現モデルの 18 種類のシンボル

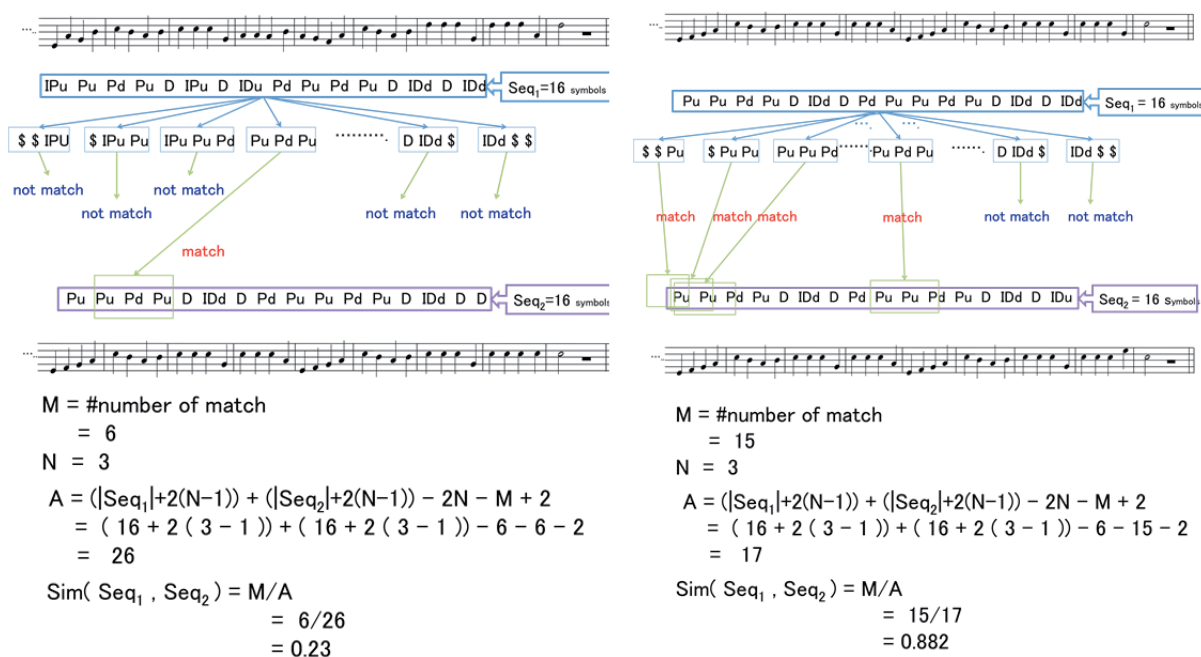


図 2, 拡張 IRM のシンボル列間の類似度計算手順

3. 楽曲間類似度

本章では楽曲間類似度計算方法について述べる。楽曲間とは拡張暗意実現モデルでの楽曲シンボル列とオリジナル暗意実現モデルでの楽曲シンボル列を指す。各シンボル列を Seq1, Seq2 として、類似度算出結果を図 2 に示す。本稿では、本計算をするにあたり python の n-gram ライブラリを使用した

4. 評価実験

4.1 評価手順

我々は Essen のフォークソングデータベースコレクション (<http://www.esac-data.org/>) から 5000 曲を用いて評価実験を行った。

5000 曲を総当たりでの比較を行い、類似度を全ての組みに対して算出した。類似度は 0~1.00 で算出され、類似度 = 0, $0.05i < \text{類似度} \leq 0.05(i+1)$ ($i=0, \dots, 18$), $0.95+0.01i < \text{類似度} \leq 0.95+0.01(i+1)$ ($i=0, \dots, 4$) のように示される 25 レンジに振り分けた。レンジごとに 2 曲で 1 組とした 5 組ずつ、合計 125 組の楽曲の組み合わせに対して主観評価実験を行った。

被験者は男女 15 人で、各組み合わせを聴き比べてもらい「どの程度似ているか」を 5 段階で評価してもらった。「とても似ている」を 5 点、「似ている」を 4 点、「どちらでもない」を 3 点、「似ていない」を 2 点、「全く似ていない」を 1 点とした。

我々は拡張暗意実現モデルとオリジナルの暗意実現モデルの両方で類似度計算を行い、被験者実験の結果を比較した。

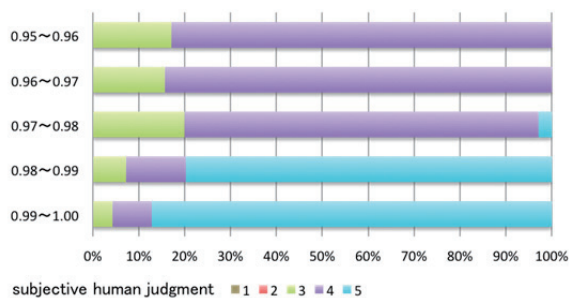
4.2 評価結果

図 3, 4, 5 に拡張暗意実現モデルを用いた場合とオリジナル暗意実現モデルを用いた場合の主観評価実験結果を示す。表 1 は類似度比較例で、赤文字になっている部分が Seq1 と Seq2 で差分が現れた部分である。

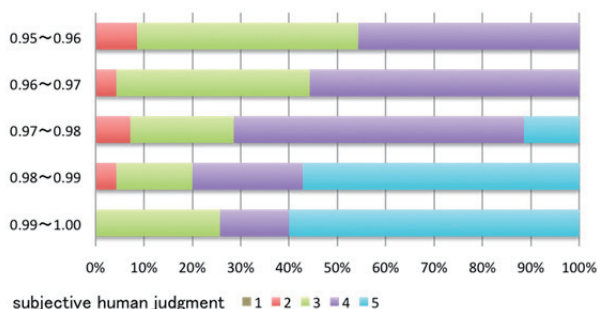
この結果より、拡張暗意実現モデルに基づき類似度計算を行った場合の方が、オリジナル暗意実現モデルに基づき類似度計算を行った場合よりも明らかに良い性能を示した。

表 1, シンボル列間の類似度計算例

$\text{Sim}(\text{Seq}_1, \text{Seq}_2)$	Extended Symbols
0.8951	Seq1: Rd IDd IDd dyad Ru Pd monad ... Ru Pd IDd Pd Pu ... Pd Pu IDu IPu Pu Vpu Ru IPd
	Seq2: Rd IDd IDd dyad Ru Pd monad ... Ru Pd Id IDd Pd Pu ... Pd Pu IDu IPu Pu Vpu Ru IPd
0.8224	Seq1: Ru Pd IDd Pd Pu IPu dyad Pu Pd IDd Pd IPu Pu dyad IPd Pd
	Seq2: Pd IDd Pd Pu IPu dyad VRu Pd Pd IDd Pd IPu Pu dyad IPd Pd
0.6941	Seq1: D IPd Pd monad IPd Pd ... IDd Pd IPd Pd IDd Pd ... IDd Pd Pu IPu Pu IPu Pd Pu IDu
	Seq2: D IPd Pd monad monad Pd ... IDd Pd IPd Pd IDd Pd ... IDd Pd Pu Pd IPu IPu Pd Pu IDu
0.6058	Seq1: ...Pu IPu dyad VPu ...Pd IDu Pd Pu ... dyad Pd D D IPu Pd IDd Pd Pu monad IPd ...
	Seq2: ...Pu IPu dyad VPu ...Pd dyad Pu Pu ... dyad Pu D D IPu Pd IDd Pd Pu monad VPd ...
0.5174	Seq1: D D monad IRd D D IRd Ru Pu D D IDd Pd ...
	Seq2: D D monad IRd D dyad D D Pu Pd Pu monad D ...

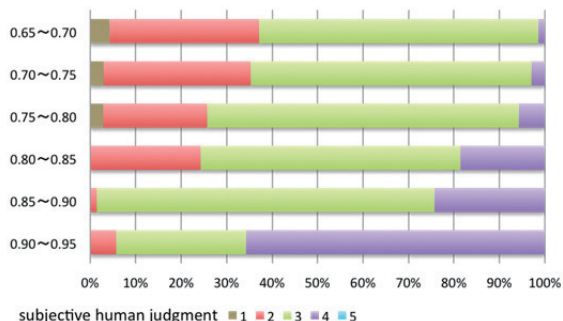


(a) 拡張暗意実現モデル

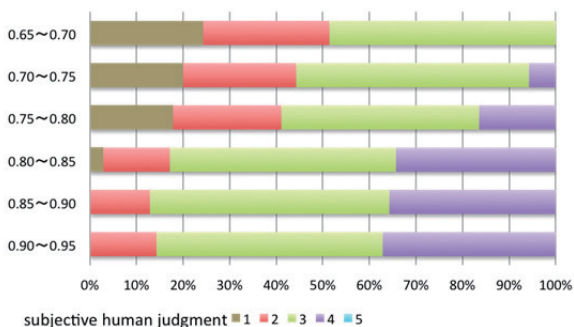


(b) オリジナル暗意実現モデル

図 3. 評価結果 (0.95<類似度≤1)



(a) 拡張暗意実現モデル



(b) オリジナル暗意実現モデル

図 4. 評価結果(0.65 <類似度 ≤ 0.95)

5. まとめ

本稿では暗意実現モデルに基づいた楽曲間類似度計算手法について提案した。提案は暗意実現モデルに拡張を行うというもので、オリジナル暗意実現モデルで定義されているシンボルに対して、音程変化を下げる/上げる方向で区別する行動の導入である。また、拡張暗意実現モデルとオリジナル暗意実現モデルでの分析を行える分析器(シンボルパーサー)を実装した。

本稿で提案した拡張暗意実現モデルによって、オリジナルの暗意実現モデルよりも適切な楽曲間類似度算出が実現できた。実際に聴き比べを行う主観的類似度評価実験の結果より、拡張暗意実現モデルによって、オリジナル暗意実現モデルと比較して精度よく類似度計算ができることを示した。

今後は、今回使用した音楽ジャンル以外の楽曲データベースを用いて、提案した類似度尺度が十分に機能するかどうかを検討する。また、広い範囲の構造を類似度計算に組み込むことでさらなる暗意実現モデルの拡張を検討する。また、現在のルールベースパーサーでは分析できない楽曲をカバーするために機械学習に基づいたパーサーの実装を行う。

参考文献

- [Narmour 1990] Eugene Narmour, "The Analysis and Cognition of Basic Melodic Structures", The university of Chicago press
- [Narmour 1992] Eugene Narmour, "The Analysis and Cognition of Melodic Complexity", The university of Chicago press
- [Hoashi 2003] Hoashi K, Matsumoto K, Inoue N., "Personalization of User Content-based Music Retrieval on Relevance Feed Back" Proceedings of ACM Multimedia , pp110-119.
- [Lampropoulos 2004] Lampropoulos A S, Sotiropoulos D N, Tsihrintzis G A, "Individualization of Music Similarity Perception via Feature Subset Selection", IEEE, International Conference on System, Man and Cybernetics 2004
- [Velardea 2013] Gissel Velardea, Tillman Weydeb, David Meredith, "An approach to melodic segmentation and classification based on filtering with the Haar-wavelet " Journal of New Music Research Volume 42, Issue 4.
- [Doraisamy 2004] Shyamala Doraisamy, Stefan Ruger "A Polyphonic Music Retrieval System Using N-Grams", Proc. of ISMIR 2004
- [Lerdahl 1983] Fred Lerdahl, "A Generative Theory of Tonal Music", The MIT Press
- [Hamanaka 2007] Msatoshi Hamanaka, Keiji Hirata, Satoshi Tojo : "FATTA: FULL AUTOMATIC TIME-SPAN TREE ANALYZER", Proceedings of the 2007 International Computer Music conference
- [Hirata 2013] Keiji Hirata, Satoshi Tojo, Masatoshi Hamanaka, "Cognitive Similarity grounded by tree distance from the analysis of K. 265/300e ", Proceedings of CMMR 2013
- [Grachten 2005] Maarten Grachten, Josep Lluís Arcos and Ramon Lopez de Mantaras, "Melody Retrieval using the Implication/Realization Model", MIREX 2012 Symbolic Melodic Similarity Results