

科学研究費助成事業 研究成果報告書

平成 27 年 6 月 19 日現在

機関番号：13901

研究種目：挑戦的萌芽研究

研究期間：2013～2014

課題番号：25540168

研究課題名(和文) 回答行列の三角化に基づく音楽類似度の個人性分析

研究課題名(英文) Analysis of individuality in subjective similarity among music songs

研究代表者

武田 一哉 (Takeda, Kazuya)

名古屋大学・情報科学研究科・教授

研究者番号：20273295

交付決定額(研究期間全体)：(直接経費) 2,900,000円

研究成果の概要(和文)：RWC 研究用音楽データベースの「ポピュラー音楽」から選択された 200 の楽曲ペアに関して 27 名の被験者が類似度を評価した。全体的な類似度とは別に、メロディ、テンポ・リズム、声質、楽器構成についての類似度も収集した。回答結果の分析から、「似ている/似ていない」の判断境界が個人毎に大きくばらつくことが示唆された。個人に最適化された楽曲間距離関数(重み付けユークリッド距離)を学習することで個人毎に主観的な楽曲間類似度を推定する実験を行った。その結果、距離関数の学習によって「声質」に関する類似性推定の精度が向上したことから、重み付けユークリッド距離を用いた個人適応の効果が明らかになった。

研究成果の概要(英文)：We conducted data collection for individuality analysis of subjective music similarity. 27 subjects evaluated similarity for 200 pairs chosen from RWC popular music database, a widely used database in the field of music information processing. Each subject also evaluated similarity for melody, tempo/rhythm, vocals and instruments for the pairs. By analyzing collected data, it is suggested that decision boundaries between similar and dissimilar pairs vary widely between individuals. Using the collected data, we trained optimized distance function between songs (weighted Euclidean distance) for each individual. In the result, training of distance function improved similarity estimation precision for vocals. From this fact, the effect of individual optimization using weighted Euclidean distance was confirmed.

研究分野：行動信号処理

キーワード：主観的類似度 回答行列

1. 研究の当初の背景

(1) ビッグデータ時代を迎え、大量のデータから知識や利用者の個性を抽出する技術が求められている。本研究では音楽楽曲間の類似度を取り上げる。

2. 研究の目的

(1) 楽曲間に「感じる」個人毎の主観的類似度を推定し、個人毎に類似楽曲の検索を行うことである。楽曲間の主観的類似度が、楽曲の音楽音響的特徴と被験者の個人性という2つの潜在要因によると考え、類似度の評価行列から2要因を抽出する手法を確立し、主観的類似度の推定方法を構築する。数量化理論類に着想を得て考案した、行列の下三角化に基づく多変量解析法を用いてこの課題に挑戦する。

3. 研究の方法

(1) 本研究では、楽曲間の主観的類似度を楽曲間類似度と聴取者の個人性という2つの要因から成ると考える。聴取者の個人性には、楽曲間のどのような特徴に着目して評価するか、及び、どのくらい似ていたら似ていると評価するという2つがあると考えられるが、先行研究より「どのくらい似ていたら似ていると評価するか」に大きな個人差があることが明らかになっていた。本研究では、楽曲間の主観的類似性が、楽曲間類似度と聴取者がどのくらい似ていたら似ていると感じるか(以下、許容度と呼ぶ)から決定されと考え、被験者による「似ている」・「似ていない」の2段階の楽曲間類似性評価データから、その2つの要因を分析する手法について考える。

(2) 本手法では、被験者×楽曲ペア(2つ

の楽曲の組)の類似性回答行列において、「似ている」要素を行列の下三角領域に集めることにより、被験者と楽曲ペアの特徴を同時に分析することができる。具体的には、被験者×楽曲の類似性回答行列の各行・各列に対し、被験者の許容度・楽曲の類似度をそれぞれ対応させ、「似ている」・「似ていない」点を共に許容度×類似度の平面上で表す。回答行列の下三角領域に「似ている」要素を集めるといふ問題を、識別問題として定式化することにより、楽曲間の類似度評価を楽曲間の類似度と被験者の許容度に分解する手法を検討する。

(3) 実験に参加した被験者は男性13名、女性14名の合計27名であった。被験者の年代は20代であった。被験者27名のうち10名が楽器経験者であった。27名という人数は、楽曲間類似度の参照データとしては必ずしも十分な量ではないが、本実験では被験者の年代を20代に限定している。被験者の年代が異なれば類似性判断の傾向も異なり、収集するデータもより多く必要になるものと考えられるが、年代を20代に限定したことによって後述する通り、統計的に有意な分析結果を得ることができた。また、20代の被験者の実験結果を用いて有用な分析手法を発見できればその手法を他の年代に適用することは可能であると考えている。

(4) 実験に使用した楽曲はRWC研究用音楽データベースの「ポピュラー音楽」[11]の楽曲100曲の内、日本のポピュラー音楽(J-Pop)スタイルによる楽曲80曲(楽曲番号: No. 1-80)を用いた。楽曲の再生区間は楽曲の一番最初のサビ区間の開始時点から30秒間とした。サビ区間の切り出しにはRWC

研究用音楽データベースへのアノテーションである AIST アノテーションを用いた。実験では、被験者に対し使用楽曲 80 曲から 2 曲を選び提示した。以下、この 2 曲を合わせて「楽曲ペア」と呼ぶ。実験において各被験者は 200 対の楽曲ペアを聴き、その各々について類似性の評価を行った。個人性を分析する目的から、全被験者が同じ 200 ペアを評価した。

(5) 収集した主観評価データを用いて、楽曲間の主観的類似度に良く対応する音響的類似度の算出手法の検討を行う。本研究で検討した、音響類似度算出手法は以下の通りである。まず、各楽曲の音響信号から短時間特徴量を抽出する。抽出した短時間特徴量を要約することにより楽曲の特徴を表す大局的な特徴ベクトルを得る。二つの特徴ベクトル間のユークリッド距離を求めることにより楽曲間の類似度を算出する。

表 1 特徴量の比較結果 (ヒストグラム)

評価の観点	特徴量セット	f_s	クラスタ数	相関
全体	TMB+ Δ + $\Delta\Delta$	16000	256	0.33
	TMB+ Δ + $\Delta\Delta$	22050	256	0.30
	TMB	16000	256	0.30
メロディ	TMB+ Δ + $\Delta\Delta$	16000	256	0.30
	TMB+ Δ + $\Delta\Delta$	22050	256	0.28
	TMB+ Δ	22050	256	0.25
テンポ・リズム	TMB+ Δ + $\Delta\Delta$	16000	256	0.30
	TMB+ Δ + $\Delta\Delta$	22050	256	0.27
	TMB+ Δ + $\Delta\Delta$	16000	512	0.27
声質	TMB	16000	256	0.16
	TMB	16000	512	0.16
	TMB	22050	256	0.15
楽器構成	TMB+ Δ + $\Delta\Delta$	16000	256	0.40
	TMB+ Δ + $\Delta\Delta$	22050	256	0.36
	TMB+ Δ + $\Delta\Delta$	16000	512	0.36

4. 研究の成果

(1) 実験の結果を表 1 に示す。表 1 は特徴量セットの中で、類似していると答えた人数との相関が最も高かった特徴量上位 3 つま

でを、構成要素毎に示している。また、表中の「特徴量セット」は特徴量セットのどれを用いたか、「 f_s 」は特徴抽出に用いた音楽データの標本化周波数 (Hz)、「クラスタ数」はベクトル量子化のクラスタ数、「相関」は似ていると評価した人数との相関をそれぞれ示している。表中の「全体」以下の 3 項目は似ていると評価した被験者数との相関が最も高かった特徴ベクトルを、「メロディ」以下の 3 項目はメロディが似ていると評価した被験者数との相関が最も高かった特徴ベクトルをそれぞれ表している。

(2) 全評価項目に対し、使用する特徴量として音色特徴量 (スペクトル重心、スペクトルロールオフ、スペクトルフラックス、高周波数エネルギー) を用い、声質以外の項目に対しては一階微分と二階微分を特徴量に加えた時に相関が高い。相関については、楽器構成に対し 0.40 と相対的に高い相関が得られた。それ以外の評価項目については類似性、メロディ、テンポ・リズムで 0.30 程度、声質では 0.16 と相関が低い。

表 2 特徴量の比較結果 (対数ヒストグラム)

評価の観点	特徴量セット	f_s	クラスタ数	相関
全体	PCA	16000	512	0.57
	PCA	44100	512	0.57
	PCA	22050	512	0.56
メロディ	INS+ Δ + $\Delta\Delta$	44100	512	0.49
	INS+ Δ + $\Delta\Delta$	44100	256	0.49
	TMB+ Δ + $\Delta\Delta$	16000	256	0.47
テンポ・リズム	INS+ Δ + $\Delta\Delta$	44100	256	0.44
	INS+ Δ + $\Delta\Delta$	22050	512	0.44
	INS+ Δ + $\Delta\Delta$	44100	512	0.43
声質	MFC+ Δ	44100	256	0.43
	MFC+ Δ	44100	512	0.43
	MFC+ Δ	44100	1024	0.42
楽器構成	ALL+ Δ + $\Delta\Delta$	22050	256	0.68
	ALL+ Δ + $\Delta\Delta$	44100	256	0.68
	PCA	44100	256	0.67

(2) 対数ヒストグラム (表 2) の場合は、

類似性と楽器構成において全特徴量に一階微分と二階微分を加えた特徴量，もしくはそれを主成分分析した特徴量，メロディとテンポ・リズムではインテンシティに一階微分と二階微分を加えた特徴量，声質では MFCC に一階微分を加えた特徴量で相関が高い．相関については類似性と楽器構成においてそれぞれ 0.57, 0.68 と相対的に高い値が得られ，その他の項目においても 0.40 以上と，ヒストグラムよりも全体的に高い値が得られている．特に，声質においてヒストグラムとの違いが顕著である．

(4) 主観的類似度と音響的類似度の対応関係は個人毎に異なると考えられるため，個人毎に最適化された音響的類似度を求めることができるかと望ましい．そこで，本研究では以下の方法により音響的類似度の個人適応を行う．算出した特徴ベクトル間の距離を重み付けユークリッド距離により算出する．重み行列 W を各々の被験者について最適化することにより，被験者毎に異なる距離関数を得ることができる．本稿では，主観評価データを用いて重み行列を最適化することにより被験者毎に最適な距離関数を得た．

(5) 実験の結果を図 1 に示す．図は評価観点毎に 5 つの条件(ユークリッド距離，他の被験者のデータを使って MLR により学習された距離，同一被験者のデータを使って MLR により学習された距離，他の被験者のデータを使って ITML により学習された距離，同一被験者のデータを使って ITML により学習された距離)で AUC を算出しその全被験者にわたる平均をとったものを示している．声質以外の評価観点においては，学習により得

られた距離関数 (MLR, ITML) と重み付けを行わないユークリッド距離とに違いがないことが分かる．この結果から，声質以外の観点では，主観的に感じる類似度と音響特徴量との対応に個人間で大きな相違が無いことが考えられる．よって，声質以外の観点では「似ている / 似ていない」の判断境界が個人性においてより重要である．

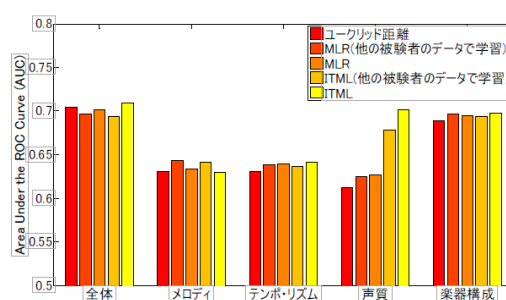


図 1：適応後の分類性能

(6) 本手法の応用として、データベースにいない新しい被験者が楽曲評価を行った際に、既存の被験者との判定パターンの類似を用いて許容度を推定し、少数の評価データから許容度の推定を行う手法を考案した。数十名分の類似性評価データを用いて、各被験者・楽曲に対する許容度と類似度をあらかじめ求めておくことで、単純な手法で許容度を推定するよりも少ないデータ数で類似性評価の予測が可能となることを示した。

5. 主な発表論文等

(1) 川淵将太, 宮島千代美, 北岡教英, 武田一哉, “楽曲間の類似判定における許容度の推定,” 情報処理学会研究報告, 2013-MUS-98-2, Mar. 2013.

(2) Shota Kawabuchi, Chiyomi Miyajima,

Norihide Kitaoka, Kazuya Takeda,
“Modeling subjective evaluation of music
similarity using tolerance,” European
Signal Processing Conference, 2013.

(3) 川淵将太, 宮島千代美, 北岡教英, 武田
一哉, “ 楽曲間主観的類似判定における個人
性分析手法の検討,” 日本音響学会 2013 年
秋季研究発表会講演論文集, 2-3-15, Sep.
2013.

6 . 研究組織

(1) 研究代表者

武田 一哉(TAKEDA, Kazuya)

名古屋大学・大学院情報科学研究科・教授

研究者番号：20273295

(2) 研究協力者

川淵 翔太 (KAWABUCHI, Shota)

名古屋大学・大学院情報科学研究科・

大学院生