

ユーザの嗜好を反映した音楽情報検索手法の検討†

2L-01

帆足 啓一郎 井ノ上 直己 (KDDI 研究所)

1. はじめに

MP3 の普及などにより、ネット上での音楽情報の流通が盛んになっており、大量の電子化音楽データの中から音楽情報を検索する音楽情報検索技術の必要性が高まっている。しかし、既存の検索システムでは曲名・アーティスト名に基づく検索や、メロディ等の音声入力により、音楽 DB の中からある特定の音楽データを検索するシステムがほとんどである [1]。本研究では、数曲の音楽データに対するユーザの評価データに基づき、ユーザが好む音楽データを検索する手法を提案する。

2. 従来手法

本研究の目標であるユーザが好む音楽データ検索に類似した研究として、Foote の研究 [2] で報告されている音楽データ自動分類があげられる。この手法では、MFCC によってパラメータ化された音楽データに対し、ツリー構造を利用したベクトル量子化 (tree-structured quantization) を行う。Foote の手法の概要を図 1 に示す。

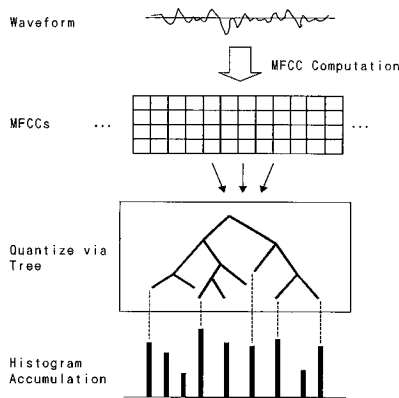


図 1 ツリーベクトル量子化手法

Foote の手法では、ツリー生成の際、対象となる音楽データに正解カテゴリが与えられた学習データを利用しており、正解カテゴリにしたがって学習データが分類されるようにツリーを生成する。次に生成されたツリーに個々の音楽データあるいはカテゴリに属する全音楽データの MFCC のフレームを入

力し、そのフレームが到達する葉 (leaf) を求める。そして、各 leaf に到達したフレームの数を算出した結果得られるヒストグラムを対象音楽データまたはカテゴリを表すベクトルとし、各カテゴリと検索対象音楽データとのベクトル類似度を算出するなどして自動分類を行う。

Foote の研究ではこの手法に基づき音楽・音声等のオーディオデータの自動分類実験を行った。また、Pye の研究 [3] では同手法に基づき、ジャンルや演奏アーティストに基づく音楽データの分類実験が報告されている。

3. 問題

Foote の手法で生成されるツリーの自動分類における有効性は、学習データの量と質に大きく影響される。前述の研究報告では、ツリー生成時に大量の学習データを利用しているうえ、各カテゴリ間の音響的特徴が大きく異なるため、自動分類が容易であると考えられる。これに対し、本研究の目標であるユーザの嗜好に基づく音楽情報検索においては、学習データとしてユーザの音楽に対する評価を収集する必要があるが、従来の研究報告で利用されている規模の学習データを収集するためにはユーザに多大な負担が発生する。また、各データの音響的特徴とそれに対するユーザの評価の関係は従来の自動分類実験で使用されたデータよりも曖昧であると考えられる。そのため、Foote の手法では、ユーザの評価に基づく音楽情報の分類・検索において十分な精度が得られない可能性がある。

4. 提案手法

本研究では、上記手法で十分な検索精度が得られない場合、ユーザからの適合フィードバック情報を利用して検索キーを更新し、検索精度を向上させる手法を提案する。本手法は、テキスト情報検索分野において幅広く利用されている検索式拡張 (query expansion) と同様の手法である。

具体的には、Foote の手法でユーザの好む音楽データを表すベクトルを生成し、検索対象音楽データとの類似度が高いデータをユーザに提示する。そのうち、ユーザが高い評価を与えたデータを再度ツリーに入力し、初期ヒストグラムにその結果を蓄積することによって新たなベクトルを作成し、同ベクトルを利用して再度検索を行うという手法である。

5. 評価実験

提案手法の有効性を検証するため、評価実験を行った。実験データとして、洋楽・邦楽のポピュラーソングからなる 107 曲を収集し、個々の曲に対し、

† Study on music information retrieval method based on user preferences.

Keiichiro HOASHI (hoashi@kddilabs.jp), Naomi INOUE
KDDI R&D Laboratories, Inc.

2名の被験者 U_1, U_2 による5段階評価(好き:5~嫌い:1)が与えられた。このデータを、評価値が4~5のデータ, 3のデータ, および1~2のデータの3つのカテゴリ(以下, C_g, C_f, C_b)に分割した。

学習の際は, 上記3カテゴリのそれぞれからランダムに選択された N 件のデータを学習データとしてツリーを生成し C_g, C_b を表すベクトルを得る。また, 上記学習データを除いたデータ集合を検索対象データとし, 同様にベクトル化した。

検索対象データ M_i とカテゴリ C_g ならびに C_b とのベクトル類似度をそれぞれ $Sim(M_i, C_g), Sim(M_i, C_b)$ とすると, 検索対象 M_i が被験者に好まれるかどうかを表すスコア S の算出方法は以下の数式により定義する。

$$S = Sim(M_i, C_g) - Sim(M_i, C_b)$$

すなわち, スコア S とは, カテゴリ C_g, C_b と検索対象データの類似度の差によって表されるものとする。スコア S を全ての検索対象データについて算出し, 検索対象データを S に基づいてソートしたデータ一覧を検索結果とする。各検索結果に対する評価基準として TREC[4]で使用されている平均精度(average precision)を利用した。

また, 適合フィードバックをシミュレートするため, 上記検索の結果得られたリストのうち C_g に属する音楽データを上位より K 件抽出し, これらのデータをツリーに入力して初期学習で得られた C_g のベクトルに蓄積する。さらに, スコア S_2 を算出する場合は, C_b に属するデータを K 件抽出し, C_b の初期ベクトルに蓄積する。この処理で生成された新たなベクトルに基づき, 再度検索対象データのスコアを算出することにより, 適合フィードバック後の検索結果を得る。なお, 本実験では学習データがランダムに選択されるため, $\{N, K\}$ のパラメータセット毎にそれぞれ12回ずつ実験を行った。

6. 結果

まず, $N=\{1,3,5,7\}$ とパラメータを設定し, 適合フィードバックを適用しない条件下での全実験の平均精度の平均値を表1に示す。なお, 比較のため検索対象データをランダムにリストアップした場合の平均精度(Random)も同表に示す。

表1 適合フィードバックなしの平均精度

Subject User	No. of training data (N)			
	1	3	5	7
U_1	0.3280	0.4237	0.3983	0.4534
U_2	0.2979	0.3578	0.3873	0.3681
Random	0.1840	0.1554	0.1255	0.1196

表1より, Footeの手法はRandomの場合を上回る精度が得られていることが明らかである一方, 被験者によって大きく精度が変わることがわかる。ま

た, 表1には被験者ごとに全実験の平均精度の平均値を計算した結果が示されているが, 個々の被験者ごとでも, 選択された学習データによって得られた平均精度が大きくばらついている。このことは, ユーザの嗜好の傾向や, 学習データに含まれる音楽データの特徴によって検索精度が大きくばらつくこと, すなわち, 少ない学習データではユーザの音楽的嗜好を正確に把握することが困難であることを示している。

次に, 適合フィードバックを適用した場合の実験結果を示す。表2および表3に, 表1の実験で得られた初期検索結果に対し, 適合フィードバック量 K を $K=\{1,3,5\}$ と設定して得られた検索結果の初期検索に対する平均精度の改善率を被験者毎に示す。

表2 適合FB後の平均精度改善率(被験者 U_1)

No. of feedback	No. of training data (N)			
	1	3	5	7
$K=1$	+13.2%	+4.0%	+10.6%	+4.5%
$=3$	+27.6%	+11.8%	+29.1%	+22.6%
$=5$	+35.8%	+21.7%	+45.0%	+28.9%

表3 適合FB後の平均精度改善率(被験者 U_2)

No. of feedback	No. of training data (N)			
	1	3	5	7
$K=1$	+19.3%	+4.5%	+5.4%	+1.7%
$=3$	+34.0%	+20.8%	+19.3%	+15.5%
$=5$	+55.4%	+35.2%	+33.6%	+29.4%

表2ならびに表3の実験結果より, 適合フィードバックを適用することにより, 全ての条件下で検索精度が向上していることが明らかである。また, 学習データが少ない条件下では特に大幅な検索精度向上が得られている。

7. まとめ

本研究では, ユーザの音楽的嗜好に基づく音楽情報検索を実現するため, ユーザからの適合フィードバックを利用した検索手法を提案し, 評価実験を通じて提案手法の有効性を証明した。

謝辞

本研究において多大な貢献をいただいたスウェーデン・Uppsala大学のErik Zeitler氏に感謝する。

参考文献

- [1] Foote: An overview of audio information retrieval, Multimedia Systems, Vol. 7 No. 1, pp. 2-11, ACM Press/Springer-Verlag, 1999.
- [2] Foote: Content-based retrieval of music and audio, Proc. of SPIE, Vol. 3229, pp. 138-147, 1997.
- [3] Pye: Content-based methods for the management of digital music, Proc. of ICASSP 2000, Vol. IV pp. 24-37, 2000.
- [4] Voorhees et al.: The 8th Text Retrieval Conference, NIST SP 500-246, 2000.