

# SVM を利用した音楽情報検索手法の検討

帆足 啓一郎<sup>†</sup> 井ノ上 直己<sup>†</sup>

KDDI 研究所<sup>†</sup>

## 1. はじめに

MP3 などの普及により、ネット上での音楽情報の流通が盛んになっており、大量の電子化音楽情報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。筆者らは、ユーザの音楽的嗜好を学習し、ユーザが好むと思われる音楽データを検索する手法について研究を進めてきた[1]。本研究では、サポートベクターマシンを利用した音楽情報検索手法を提案し、評価を行う。

## 2. 従来手法

本研究では、Foote の研究[2]で提案されたツリー構造を利用したベクトル量子化手法（以下、TreeQ）を利用し、検索対象音楽データのベクトル量子化を行う。TreeQ の概要を図 1 に示す。

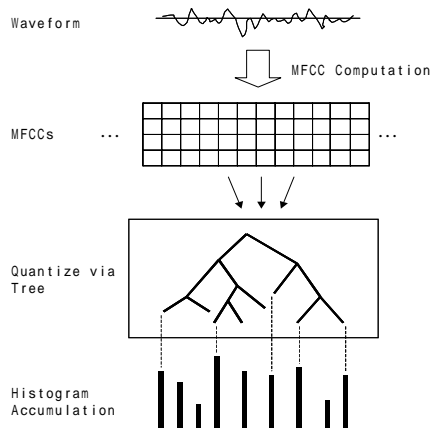


図 1 ツリーベクトル量子化手法概要

TreeQ 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する。筆者らは、ユーザの好きな曲・嫌いな曲を学習データとして利用し、好きな曲ならびに嫌いな曲のベクトルと検索対象音楽データとのベクトル類似度を算出す

<sup>†</sup>Study on music information retrieval method based on SVM.

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

ることによりユーザが好むと思われる曲を検索する手法を提案し、評価を行った[1]。

## 3. SVM を利用した音楽情報検索手法

サポートベクターマシン(SVM)は、ベクトル空間に分布する正負例を分離する多くの超平面の中から最適な分離超平面を求めるアルゴリズムであり、テキスト自動分類をはじめとする分類タスクにおいて高い性能を示している[3]。SVM の概要を図 2 に示す。

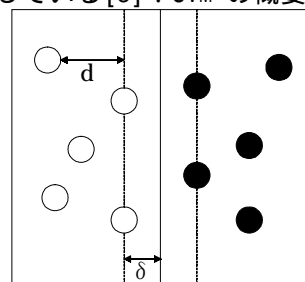


図 2 SVM 概要

SVM の利点の1つとして、学習データが少量であっても高い精度での分類が可能である点があげられている。ユーザに嗜好に基づく音楽情報検索システムを実現する場合、ユーザから事前に音楽嗜好に関する情報が大量に提供されるとは考えにくい。筆者らの研究でも少量の学習データに基づいた音楽情報検索手法の有効性が確認されたが、本研究では SVM を利用した手法を提案し、従来手法との比較を行う。

## 4. 評価実験

### 4.1. 使用実験データ

本実験では、HMV Japan ウェブサイト (<http://www.hmv.co.jp/>) 2001 年上期(1月~6月)の週間チャートのトップ 10 に登場した全 CD の中から、ベスト版およびコンピレーション版を除いた 60 枚の CD に収録されている全ての楽曲(756 曲)を実験データとして使用した。さらに、ユーザの音楽嗜好のデータとして、実験データに含まれる全ての楽曲に対し、被験者 12 名が各々の音楽嗜好に基づく 5 段階評価(1:嫌い~5:好き)を与えたユーザ評価データを収集した。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。

### 4.2. 実験手法

学習の際は、前節の 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する。ベクトル類似度を利用した手法の場合は、 $C_g, C_b$  を表すベクトルを得、これらのベクトルと検索対象データとの類似度を算出する。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。

一方、SVM に基づく手法の場合は、学習の過程で得られたツリーを利用し、学習データに含まれる個々の楽曲をベクトル化し、これらのデータに基づき SVM の分離超平面を算出する。検索対象データについては、分離超平面の正例側からの距離（図 2 内の  $d$ ）を算出し、その長さを基準にソートしたデータ一覧を検索結果とする。

なお、学習データのランダム性を考慮し、本実験では各  $N$  の設定値に対し、10 回ずつ実験を行う。

#### 4.3. 結果

表 1 に、ベクトル類似度 (Vec) ならびに SVM を利用した音楽情報検索手法で得られた検索結果に対する上位 5 件および 10 件での精度 (Prec@5, Prec@10) の平均値を示す。また、比較のために全検索対象データをランダムに並び替えた場合の精度も示す (Rnd)。なお、本実験では、 $C_g$  に含まれるデータを適合データとして精度を算出した。

表 1 ベクトル類似度ならびに SVM の検索精度

N	Prec@5		Prec@10	
	Vec	SVM	Vec	SVM
1	0.5183	0.4367	0.4867	0.4700
2	0.5217	0.4583	0.4992	0.4658
3	0.5533	0.4967	0.5358	0.5167
4	0.5217	0.5217	0.5042	0.5175
5	0.5633	0.5700	0.5650	0.5433
Rnd	0.4197		0.4187	

表 1 に示された結果より、両手法ともランダムを上回る精度が得られ、有効性が確認された。また、学習データ ( $N$ ) が増加するほど、SVM の検索精度が向上している一方、ベクトル類似度の場合は学習データ量と検索精度の間に相関が見られず、全体的には SVM を上回る精度が得られている。

#### 5. 適合フィードバック実験

前節の実験結果から、少ない学習データでは SVM で

の検索精度が低いことがわかった。そこで、情報検索で広く利用されている適合フィードバック (FB) 技術を利用し、精度向上を図る。具体的には、上記の検索の結果、スコア上位  $M$  件のデータに対しユーザが FB 情報を与える状況を想定し、学習データにこの  $M$  件を加えて分離超平面を再算出した。

表 2 に、適合 FB 後の SVM の検索精度 (Prec@10) を示す。なお、適合 FB 前の検索結果との比較のため、初期検索の結果から FB に利用した  $M$  件のデータを除いて算出された検索精度も示す (カッコ内)。

表 2 適合 FB 前後の検索精度 (Prec@10)

	$M=3$	$M=5$	$M=7$
$N=1$	0.4658 (0.4783)	0.4817 (0.4633)	0.4517 (0.4517)
2	0.4767 (0.4783)	0.4458 (0.4608)	0.4758 (0.4550)
3	0.5142 (0.5067)	0.5083 (0.4883)	0.5133 (0.4767)
4	0.5058 (0.4875)	0.5083 (0.4875)	0.4842 (0.4783)
5	0.5442 (0.4967)	0.5158 (0.4950)	0.5342 (0.5000)

表 2 より、適合 FB の結果、学習データが多い条件下では検索精度が向上していることがわかる。しかし、学習データおよび適合 FB データが少ない条件下では、検索精度の向上が見られない。

#### 6. 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タスクに対し、SVM を利用した検索手法を提案し、従来手法との比較を行った。評価実験の結果、適合 FB を利用することにより検索精度向上が見られたが、従来手法と比べ検索精度が劣ることが明らかになった。今後の課題として、SVM ならびに適合 FB において最適な音楽データの抽出方法などについて検討する必要があると考えられる。また、SVM に対する適合 FB 手法そのものについても再検証の必要がある。

#### 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 Nina Ewerlof 氏に感謝する。

#### 参考文献

- [1] 帆足, 井ノ上: ユーザの嗜好を反映した音楽情報検索手法の検討, 第 64 回情処全大, Vol 2, pp 11-12, 2002.
- [2] Foote: Content-based retrieval of music and audio, Proc. of SPIE, Vol 3229, pp 138-147, 1997.
- [3] Vapnik: Statistical learning theory, A Wiley-Interscience Publication, 1998.