

複数ユーザの音楽嗜好特徴空間を用いた楽曲推薦システム

永田 裕二[†] 帆足 啓一郎[‡] 菅谷 史昭[‡] 甲藤 二郎[†]

早稲田大学理工学部[†] KDDI 研究所[‡]

1. はじめに

近年、携帯型音楽再生機やオーディオデータフォーマットの急速な普及に伴い、膨大な楽曲を蓄積し再生することが可能になっている。このような背景を元に、ユーザにとって本当に必要な楽曲(ユーザの好む楽曲)を推薦するシステムの需要が高まっている。

本研究は、大量の音楽データの中から未知の音楽情報を発見するユーザ嗜好に基づいた content-based 音楽検索システムにおいて、より幅広く楽曲を推薦するために、複数のユーザの音楽嗜好特徴空間を用いて楽曲を推薦する手法を提案する。

2. ユーザ嗜好に基づく音楽情報検索手法

2.1 従来手法

従来の音楽検索システムは、帆足らの手法[1]により、ユーザ嗜好に基づく音楽情報検索を行う。具体的には、特定の特徴空間上において、ユーザの主観評価に基づき、ユーザプロフィールを作成する。ユーザプロフィールと検索対象音楽データをコサイン距離により類似度を算出し、高いものを推薦する(ベクトル類似度手法)。尚、ユーザプロフィールも検索対象音楽データもツリーベクトル量子化手法(TreeQ)[2]を用いてベクトル化されている。さらに、推薦された楽曲に対するユーザの評価を適合フィードバック(relevance feedback)情報として利用し、ユーザプロフィールを更新する。

2.2 課題

従来手法で用いられている適合フィードバック手法は、対話的に検索精度を改善するという有効性が示されている[1]。一方、音楽の嗜好情報は曖昧性を持っており、特徴空間上の離れた場

所にある楽曲を好きなユーザ(複数解を持つユーザ)が存在する。適合フィードバック手法は、1つの最適解に到達するように解空間を移動するものであるため、複数解をもつユーザには、そのうちのいずれか1つの局所解に収束し、同じような曲ばかりが検索されてしまう。つまり、ユーザの好む楽曲の一部しか推薦出来ないことがある。

3. 提案手法

上記の問題を解決するために、本手法では、複数のユーザの嗜好情報に基づいてそれぞれの特徴空間を作成し、他ユーザの特徴空間上にユーザプロフィールをマッピングし、ベクトル類似度手法を用いて楽曲推薦を行う手法を提案する。図1に楽曲推薦方法の概念図を示す。推薦対象であるUser Aの嗜好情報をもとに、ユーザの特徴空間上でユーザプロフィールを作成する。次に、その際に用いたAの嗜好情報を学習データとしてUser Bの特徴空間上でユーザプロフィールを作成する。そして、作成されたユーザプロフィールと類似度の高い楽曲をAに推薦する。この処理を、推薦対象を変更し、複数のユーザで同様の操作を繰り返し行う手法である。この手法により、従来手法と比較して、より幅広い楽曲が推薦されることが期待される。

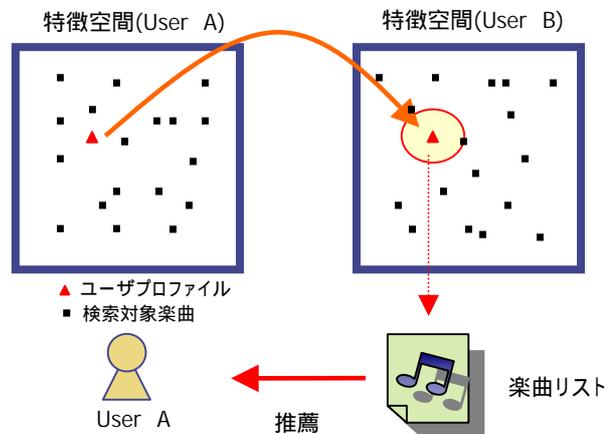


図1 提案手法(概念図)

Music Recommendation System Based on Feature Space of Multiple Users.

[†]Yuji Nagata [‡]Keiichiro Hoashi [‡]Fumiaki Sugaya [†]Jiro Katto

[†]Department school of Science & Engineering, Waseda University [‡]KDDI R&D Laboratories inc.

4. 実験

前述の提案手法の有効性を検証するために、以下の評価実験を行う。本実験に際して、検索対象楽曲 6863 曲に対し、被験者 6 名が主観評価を行ったデータを使用する。各被験者の評価結果の一覧を表 1 に示す。

表 1 楽曲評価データ

User	Good	Soso	Bad	Total
A	317	370	529	1216
B	552	490	579	1621
C	691	1148	673	2512
D	198	271	278	747
E	69	283	338	690
F	1934	528	554	3016

はじめに、上記の評価データにおいて、各ユーザの「Good」「Bad」と評価した楽曲を 10 曲ずつ選曲し、ユーザごとに特徴空間を作成しておく。次に、特定の推薦対象のユーザにおいて、「Good」「Bad」と評価した楽曲を 10 曲ずつ選曲し、学習データとする。その学習データに基づき、各特徴空間上でそれぞれユーザプロフィールを作成し、類似度を算出することで、推薦楽曲リストをそれぞれ作成する。

提案方式による楽曲推薦結果の多様性を評価するため、ユーザと他のユーザの特徴空間上で推薦された楽曲との上位 100 曲に対して、それぞれ重複率を算出することで比較を行う。ここでは、推薦対象のユーザを変更し、被験者すべてに対して実験を行った。また、ユーザごとに学習データを 3 回変更し、ユーザプロフィール作成以降の操作を繰り返すことで平均化を行った。推薦曲リストの重複率の結果を表 2 に示す。

表 2 推薦曲リスト間重複率

		特徴空間(feature space)					
		A	B	C	D	E	F
User	A		0.3%	2.3%	0.0%	2.3%	4.7%
	B	0.0%		0.3%	5.3%	0.3%	2.3%
	C	5.7%	0.0%		24.7%	24.0%	0.0%
	D	0.0%	17.0%	6.0%		5.3%	2.7%
	E	8.0%	5.7%	12.0%	12.0%		9.3%
	F	10.3%	0.7%	0.3%	0.0%	5.0%	

表 2 の結果を見ると、他ユーザの特徴空間上で検索された楽曲とユーザの特徴空間上で検索

された楽曲は平均 5.6%の重複率しか存在しなかった。つまり、94.4%は別の楽曲が推薦されていることになり、幅広く楽曲が推薦されたとみなすことができる。

表 3 検索精度(User C)

特徴空間	Good	Soso	Bad
A	46.2%	40.5%	13.3%
B	39.3%	48.1%	12.6%
C	36.3%	55.2%	8.5%
D	46.5%	39.0%	14.5%
E	39.2%	46.2%	14.6%
F	32.4%	51.9%	15.7%

次に、推薦された楽曲がユーザの嗜好に適合しているかどうかを検証するために、評価数が多く、データ的にもバランスの取れている User C のデータに注目して検索精度の評価を行った。検索精度結果を表 3 に示す。この際に用いた学習データは、表 2 で使用したものと同一である。評価方法については、各特徴空間において推薦された楽曲の上位 100 曲について、表 1 に存在した評価データの中でどのように評価していたかを検索精度とした。表 3 の結果によると、ユーザ自身の特徴空間と他ユーザの特徴空間で検索された検索精度とは、精度が上がるものと下がるもの両方の結果が得られた。この結果より、ユーザによって検索精度向上に適した特徴空間があると考えられる。今後、ユーザ間の特徴空間の相関が検索精度に与える影響について調べる必要がある。

5. まとめ

本研究では、ユーザの音楽嗜好に関する情報検索手法において、複数のユーザ嗜好を用いた音楽情報検索手法を提案した。推薦曲リストの重複率、検索精度の評価実験を行い、局所解に収束したユーザに対する有効な解決法の 1 つであることが検証された。

6. 参考文献

- [1] 帆足, 井ノ上: "ユーザの音楽嗜好に基づく音楽情報検索手法", 情処研報 SIGMUS49-14, pp79-84, 2003.
- [2] Foote: "Content-based retrieval of music and audio" Proceedings of SPIE, Vol 3229, pp 138-147, 1997.