

# ソーシャルタグをクラスラベルとする距離学習を用いた類似楽曲検索システム

Content-based music information retrieval by distance metric learning using SNS tags

萩原 智彰<sup>\*1</sup> 山田 誠二<sup>\*1\*2</sup>  
Tomoaki Hagihara Seiji Yamada

<sup>\*1</sup>東京工業大学大学院 <sup>\*2</sup>国立情報学研究所／総合研究大学院大学  
Tokyo Institute of Technology National Institute of Informatics / SOKENDAI

Recently, increasing digital music collections make us difficult to find new favorite music which we are really interested in. To overcome this problem, content-based music information retrieval has been studied. It enables users to search similar music with users choose, however few method can improve search result interactively. In this paper, we propose a new method that allows users to reflect their search intention. By selectiong SNS tags, users can naturally inform their intention to a music retrieval system, and the system utilize user feedback to modify a similarity function. we employ distance metric learning to utilize tag information to modify a feature vector space.

## 1. 研究背景

近年、定額楽曲配信サービスの登場などにより、個人で取り扱える楽曲の量は爆発的に増えているが、ユーザが自分の好む楽曲を見つけることはますます困難になっている。ユーザの楽曲発見を支援するためにいくつかの技術が研究されている。中でももっとも一般的な手法はユーザの購買履歴を元に推薦を行うものである。このような手法は協調フィルタリングと呼ばれ、amazon.comなどのサービスで実際に利用されている[神畷 2007]。しかしながら協調フィルタリングをベースとした推薦システムには、購買履歴が十分な量集まらないと適切な推薦が行えないなどいくつかの問題点が指摘されている[神畷 2007]。また、コンテンツの内容を一切考慮せずに推薦楽曲を決定するため、楽曲の特性が推薦結果に反映されないという問題がある。

一方、楽曲の音響信号を利用してユーザの楽曲発見を支援する手法として類似楽曲検索が挙げられる[Casey 2008]。この技術を用いることで、ユーザは楽曲を指定してその楽曲に類似した曲を獲得することができる。類似楽曲検索では、楽曲の音響信号をもとに複数の特徴量を抽出し、特徴ベクトルを作成する。楽曲と楽曲の類似の度合いは、特徴ベクトル同士の距離を事前に定めた類似度関数に基づいて定義する。しかしながらユーザの検索意図は多様であるため、あるひとつの類似度関数を利用するだけでは、ユーザの求める検索結果が得られない場合がある[小林 2005]。そのため、類似度関数を各ユーザごとにパーソナライズする手法が求められる。

本研究では、SNS上のタグをクラスラベルとした距離学習手法を利用しインタラクティブに改善を行う類似楽曲検索手法を提案する。提案手法により、ユーザは表示されたタグの中から自分の検索意図を反映したタグを選択することによって検索結果の改善を行うことが可能となる。

## 2. 関連研究

検索結果をインタラクティブに改善することを可能にする代表的な手法として適合性フィードバックが挙げられる。適合性フィードバックではユーザに検索結果に対して適合非適合の情

報を返し、システムはこのユーザフィードバックをもとに次式により検索クエリを修正する[Christopher 2012]。

$$\vec{q}_m = \alpha \vec{q}_0 + \beta \frac{1}{|D_r|} \sum_{\vec{d}_j \in D_r} \vec{d}_j - \gamma \frac{1}{|D_{nr}|} \sum_{\vec{d}_j \in D_{nr}} \vec{d}_j \quad (1)$$

ここで $\vec{q}_0$ が元のクエリベクトルであり、 $D_r$ 、 $D_{nr}$ は適合アイテムと非適合アイテムの集合である。 $\alpha$ 、 $\beta$ 、 $\gamma$ は重みパラメータをあらわす。クエリベクトルは $\vec{q}_0$ から出発し、適合アイテムの重心方向に向かって、非適合アイテムの重心から遠ざかる方向に、ある程度の距離移動する。この操作により得られた新しい検索クエリ $\vec{q}_m$ を用いて再度検索することにより検索結果の改善を行うことができる。文書検索の分野において、多くの場合、文書及び検索クエリはbag-of-wordsの特徴ベクトルで表現される。このとき検索クエリに含まれる単語数は少数となるため、検索クエリの特徴ベクトルは非常に制限された表現となる。このため、検索クエリと文書のベクトル表現に大きな差異が生まれ、質の悪い検索結果の原因となる。適合性フィードバックは、ユーザの指定した文書に検索クエリを近づける方向に修正することでこの問題に対処する手法である。しかしながら、本研究のような類似楽曲検索の場合、検索クエリは楽曲そのものとなるため、検索結果として得られる楽曲と検索クエリは特徴ベクトル空間上でお互いに近い位置に配置される。このとき適合性フィードバック手法を利用して検索クエリベクトルの修正を行っても、検索クエリと楽曲との間の差異が小さいためクエリの修正量は小さくなり、検索結果の大きな改善にはつながらないものと考えられる。

検索結果の改善を行うほかの手法としてランキング学習手法がある[Li 2009]。ランキング手法もまた、検索結果に対して適合非適合または選好順序を付け、それをもとに検索結果の改善を行う。しかしながら適合性フィードバックもランキング学習も、検索により得られた集合に対して評価させる手法であるため、得られたアイテムの間にユーザから見て大きな差異がない場合、検索結果に対して適切なフィードバックを送ることが困難になると考えられる。

齊藤らの研究[齊藤 2011]では、楽曲の集合を可視化した上でユーザに表示し、適合する楽曲を選択してもらい対話的遺伝的アルゴリズム手法を利用して次に表示する楽曲の選択を行っ

ていくことでユーザーの新たな楽曲発見の支援を試みている。また、類似楽曲検索に対して検索結果の個人化を試みた研究として小林らの研究 [小林 2005] があげられる。小林らは楽曲と楽曲の類似度をマハラノビス距離の形式で定義し、行列の値をパラメータとし、非適合とされる楽曲が適合とされる楽曲よりも類似度が低くなるよう学習を行うことにより改善を行う。しかしながらこれらの手法でも何らかの方法で適合非適合の情報を獲得する必要があるため、上述した問題点は解決されていない。

これらの手法に対して、我々の提案する手法では、それぞれの楽曲を評価することなく、検索意図に対応したタグを選択するという、ユーザーが自然に意図を伝える方法を通して検索結果の修正を行えるという利点がある。次章で提案手法の手続きについて詳しく解説する。

### 3. 提案手法

#### 3.1 概要

提案手法の流れを図 1 に示す。

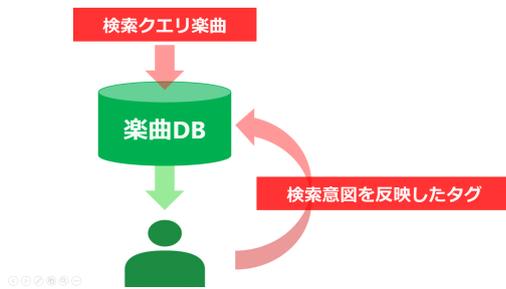


図 1 処理の流れ

ユーザーに楽曲が与えられると、まず、システムはあらかじめ定められた類似度関数を元に検索結果を返す。その際、検索して得られた楽曲に含まれるタグを同時に表示する。ユーザーは自身の重視する要素を表現しているタグを選択し、再検索することで検索結果の改善を行うことができる。

#### 3.2 距離学習アルゴリズム

類似度関数の修正を行うアルゴリズムについて記述する。学習の流れを図 2 に示す。

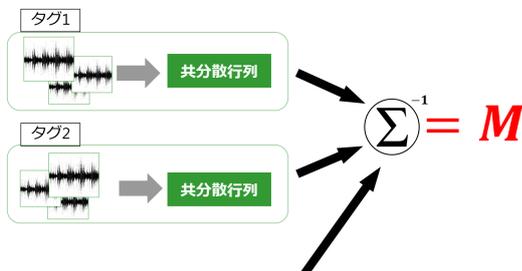


図 2 距離学習

本手法では楽曲同士の類似度を次式 (2) のマハラノビス距離で定義する。

$$d(f_i, f_j) = (f_i - f_j)^T M (f_i - f_j) \quad (2)$$

ここで  $f_i, f_j$  はそれぞれ  $i, j$  番目の曲の特徴ベクトルを表している。 $M$  はマハラノビス行列であり、初期値として単位行列が与えられる。 $M$  が単位行列であるとき式 (2) はユークリッド距離と一致する。 $M$  の適切な値を学習するための手法はいくつか提案されているが、本研究では学習が高速に行える教師あり距離学習手法である Relevant Component Analysis(RCA)[Bar 2005] を用いる予定である。RCA はクラスラベルがつけられたデータをもとに学習するアルゴリズムである。全ラベル付データ数を  $p$ 、クラス数を  $k$ 、クラス  $j$  に含まれるデータ数を  $n_j$ 、各クラスの平均ベクトルを  $m_j$  とすると  $M$  は式 (3)(4) のように修正される。

$$C = \frac{1}{p} \sum_{j=1}^k \sum_{i=1}^{n_j} (f_{ji} - m_j)(f_{ji} - m_j)^T \quad (3)$$

$$M = C^{-1} \quad (4)$$

ここで、 $f_{ji}$  は  $j$  番目のクラスの  $i$  番目の楽曲と対応した特徴ベクトルを表している。RCA では行列  $M$  は、クラス内共分散の大きい要素は重みが小さく、共分散の小さい要素は重みが大きくなるよう修正される。本手法ではユーザーが選択したタグをひとつのクラスラベルとして考え、各クラスごとに対応するタグを含む楽曲の平均ベクトルを算出し、上述のアルゴリズムによって類似度関数の学習を行う。

### 4. 評価実験

ユーザーに提示したタグをもとに類似度関数を修正することにより、検索結果が改善されることを確認するため、参加者実験を行うことを計画している。参加者実験の詳細について図 3 および下記に示す。

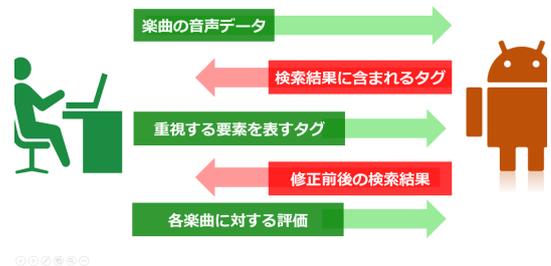


図 3 評価実験

まず、参加者は自身の好む楽曲をシステムに対し入力する。システムはその楽曲に似た曲を順に  $n$  個取り出し、これを  $s_1$  とする。次に  $s_1$  の楽曲に付いたタグを参加者に表示する。このときどのような楽曲が検索により得られたかはユーザーに表示しない。参加者はそれらのタグの中から重要と考える要素を反映した表現をもつタグを 1 つ以上選択する。システムはそれらのタグをもとに前述の手法によって類似度関数を修正し、修正後の関数に基づいて、参加者に指定された楽曲ともっとも類似している楽曲を  $n$  個取り出す。これを  $s_2$  とする。修正前後の二つの関数より得られた楽曲  $s_1, s_2$  の和集合を順序をランダムにして表示し、それぞれの曲についてどの程度自分の好みに適合するかを評価してもらう。評価値は 0 から 1 の間の得点を 0.1 点刻みで付けてもらう予定である。得られた評価を元に修正前後の評価結果に対して式 (5)(6) に示すラン

キング指標である  $nDCG$ (normalized Discount Cumulative Gain)[Jrvelin 2000] を算出する。

$$DCG_m = \sum_{l=1}^m \frac{2^{y_l} - 1}{\log_2(l+1)} \quad (5)$$

$$nDCG_m = \frac{DCG_m}{\max DCG_m} \quad (6)$$

ここで  $DCG_m$  はランキングの  $m$  番目までの評価値,  $y_l$  は順位が  $l$  番目である楽曲に付けられた評価値を表す.  $\max DCG_m$  は  $m$  番目までの理想的なランキングの  $DCG_m$  評価値を表しており,  $nDCG_m$  の値は 0 から 1 までの間の値に正規化される.  $nDCG$  は情報検索の分野で用いられる指標で検索結果のランキングが理想的なランキングにどれぐらい近いかをあらわす指標である.  $nDCG$  には, ユーザのアイテムに対する選好順序の情報を必要とせず, 各アイテムについての評価さえ得られれば算出できるというメリットがある. 本研究では, 楽曲集合に対して選好の順序を記述させるよりも, 各楽曲をそれぞれ独立に評価させる方がユーザにとって負担が少ないと考え,  $nDCG$  を採用した.

## 5. 実験に用いるデータ

本研究の実験は参加者による評価を含む実験であるため, 楽曲のタグ, 特徴ベクトルのほかに, 楽曲そのものが必要となる. そのようなデータセットは筆者の知る限りでは存在しない. そこで, 我々は本実験のために last.fm<sup>\*1</sup>, 7digital<sup>\*2</sup> の二つの API を利用してデータセットの構築を行う. last.fm は音楽に特化した SNS であり, サービスを通して得られた楽曲のメタデータを公開している. 本実験では API を利用し, もっとも人気のある楽曲 3000 曲のメタデータを獲得した. このメタデータ内にはタイトル, アーティスト名などの一般的な情報のほかに各楽曲につけられたタグの情報を含んでいる. データに含まれる楽曲の再生可能な音響信号情報を獲得するために 7digital で公開されている API を利用する. 7digital API を利用することで各楽曲ごとに 30 秒の長さの試聴用データを入手できる. 本実験ではこの試聴用データをもとに, 特徴抽出および参加者実験を行う予定である. 特徴抽出は先行研究[斉藤 2011] で用いられている MIRtoolbox<sup>\*3</sup> を用いることを計画している.

## 6. まとめと今後の課題

本論文では, 類似楽曲検索において, ユーザがインタラクティブに検索結果の改善を行える方法を考案した. 提案手法では, 検索結果改善のために, ユーザはタグの選択を行い, システムはそのタグを含む楽曲より類似度関数の修正を行い, 検索結果を修正する. さらに, 提案手法により, 検索結果が改善されることを確認するための参加者実験を設計した. その際に必要となる要素を明らかにし, その獲得方法を示した. 今後は提案するシステムを実装し, 参加者実験を行う予定である.

## 参考文献

- [神寫 2007] 神寫 敏弘: 推薦システムのアルゴリズム (1), 人工知能学会誌 22 巻 6 号 (2007)
- [Casey 2008] Casey, Michael A., et al.: Content-based music information retrieval: Current directions and future challenges, Proceedings of the IEEE 96.4 668-696 (2008)
- [Christopher 2012] Christopher D. Manning Prabhakar Raghavan Hinrich Schütze: Introduction to Information Retrieval(2008) (邦訳: 岩野和生, 黒川利明, 濱田誠司, 村上明子 情報検索の基礎 p157-p172 (2012))
- [Li 2009] Li, Hang.: Learning to rank, Tutorial given at ACL-IJCNLP, August. research. microsoft.com/enus/people/hangli/li-acl-ijcnlp-2009-tutorial.pdf (2009).
- [小林 2005] 小林和也, 長谷山美紀: 個人の嗜好を考慮した楽曲間の距離尺度とその類似楽曲検索への応用に関する考察, 社団法人映像情報メディア学会技術報告 ITE Technical Report Vol.33 No.44 PP3-6 (2005)
- [Bar 2005] Bar-Hillel, A., Hertz, T., Shental, N., Weinshall, D.: Learning a mahalanobis metric from equivalence constraints, Journal of Machine Learning Research, 6(6), 937-965. (2005)
- [Jrvelin 2000] Jrvelin, K., Kekilinen, J.: IR evaluation methods for retrieving highly relevant documents. In Proceedings of the 23rd annual international ACM SIGIR conference on Research and development in information retrieval (pp. 41-48). ACM. (2000)
- [斉藤 2011] 斉藤 優理 伊藤 貴之: 特徴量空間における対話型進化計算を用いた楽曲提示インタフェース, 2011 年度人工知能学会全国大会 (2011)

\*1 The Last.fm Dataset — Million Song Dataset <http://labrosa.ee.columbia.edu/millionsong/lastfm>

\*2 7digital Developers <http://developer.7digital.com/>

\*3 MIRtoolbox <https://www.jyu.fi/hum/laitokset/musiikki/en/research/coe/materials/mirtoolbox>