

HMMを用いた楽曲推薦システムの構築

Development of a Music Recommendation System Using HMM

鹿又広行*

小林利彰*

東海林智也*

Hiroyuki Kanomata

Toshiaki Kobayashi

Tomoya Tokairin

1 はじめに

近年、音楽再生機能を持つ携帯電話やポータブル音楽プレイヤー等の普及によって楽曲のインターネット配信は生活に密着したものになりつつある。また聴取者の楽曲に対する嗜好(印象)が多様化しているため、個別の聴取者が好む楽曲を検索する楽曲推薦システムの研究も盛んにおこなわれている [1][2][3][4][5]。

従来の推薦手法では、楽曲全体を通した聴取者の印象や他者の嗜好情報をもとにして楽曲の推薦をおこなっているものが多い。しかし、聴取者の楽曲に対する印象は楽曲の時間進行によって逐次変動するため、聴取者が常に楽曲に対して同じ印象を持ち続けること無いと思われる。さらにこの印象変化は聴取者ごとに異なると考えられるため、個別の聴取者の楽曲に対する印象変化の情報をシステムに組み込むことで効果的な楽曲推薦システムを構築できると思われる。

そこで本研究では、楽曲のコード進行と聴取者の印象変化を HMM (Hidden Markov Model: 隠れマルコフモデル) によりモデル化して楽曲推薦システムを構築する。構築するシステムは、ある楽曲に対する聴取者の印象変化を推定し、推定した印象変化からその楽曲の推薦値を計算し、適切に定めたしきい値よりも推薦値が大きい楽曲を聴取者に推薦する。

2 楽曲のコード進行と聴取者の印象変化の時系列データ化

はじめに楽曲のコード進行と聴取者の楽曲に対する印象変化を時系列データに変換する [6]。

コード進行に関しては、楽曲に含まれるコードを楽曲の先頭から 2 つを組にして取り出し、それらの組の集合をコード進行情報とする。なおメジャーコード、マイナーコードなどの違いを考慮すると状態数が膨大になるため今回は考慮しないことにする。さらに調の違いも考慮して、トニックのルート音と同じルート音を持つコード全てに状態番号 0 を割り当て、トニックのルート音から k だけ半音上の音をルート音とするコード全てに状態番号 k を割り当ててコード進行情報を数値化する。最後に、時刻 t 、($1 \leq t \leq T$) におけるコード進行情報を c_t とおくと全体のコード進行は $C^T = c_1 c_2 \cdots c_T$ という時系列データで表される。例えば、ある楽曲のコードが先頭から C F G C F C の順で現れるとき、 $\{\{C, F\}, \{G, C\}, \{F, C\}\}$ がそ

の楽曲のコード進行情報となり、C メジャースケールのもとでは $\{\{0, 5\}, \{7, 0\}, \{5, 0\}\}$ と数値化される。従って全体のコード進行の時系列データは $C^3 = c_1 c_2 c_3$ 、 $c_1 = \{0, 5\}$ 、 $c_2 = \{7, 0\}$ 、 $c_3 = \{5, 0\}$ となる。

また印象変化情報に関しては、状態(印象)数が M 個の場合には各状態に 0 から $M - 1$ 番までの数値を割り当てる。本研究では嗜好(好き、嫌い)だけを扱うため、単純に「好き」(状態番号 1)、「嫌い」(状態番号 0)の $M = 2$ 状態とする。さらにコード進行の時系列データと同様に、時刻 t 、($1 \leq t \leq T$) における印象を s_t とおくと、印象変化は $S^T = s_1 s_2 \cdots s_T$ という時系列データの形式で表わされる。

3 コード進行と印象変化の HMM によるモデル化

あるひとつの楽曲に対する聴取者のコード進行データと印象変化データの時系列の集合を $\{C^T, S^T\}$ とする。ここで、 $\{C^T, S^T\}$ の同時確率を $P(C^T, S^T)$ とすると、 S^T が未知、かつ C^T が既知であるときに、この同時確率を最大にする \hat{S}^T が S^T の最尤推定量となる。

本研究では \hat{S}^T を求めるために HMM を適用する [6]。 S^T を HMM の(未観測である)内部状態、 C^T を実際に観測されたシンボル出力として、図 1 の様に HMM を構成する。ここで a_{ij} は内部状態 i から j に遷移する確率、 $b_i(k)$ は内部状態 i の時にシンボル k が出現する確率である。

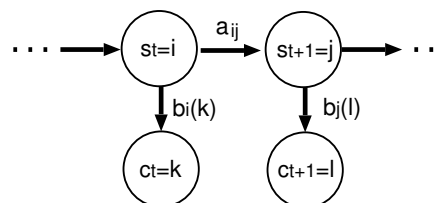


図 1: HMM によるコード進行と印象変化のモデル化

4 楽曲推薦システムの構築

図 1 において、 a_{ij} と $b_i(k)$ は実際には未知であるので、対象となる聴取者に対していくつかの楽曲の印象変化をサンプリングし、 a_{ij} と $b_i(k)$ の最尤推定量を求めて a_{ij} と $b_i(k)$ の代わりに使用する。なお、楽曲全体に対する嗜好(好き、嫌い)も同時にサンプリングしておく。

*函館工業高等専門学校情報工学科

次に楽曲のコード進行データ C^T から，印象変化の推定値 \hat{S}^T を viterbi アルゴリズム [7] などを用いて求めて実際の S^T と比較する．ここで評価基準としては 2 乗誤差基準

$$r(S^T, \hat{S}^T) = \frac{1}{T} \sum_{t=1}^T (s_t - \hat{s}_t)^2 \quad (1)$$

を使用する．なお本研究では状態数として好き (1)，嫌い (0) の 2 通りのみを使用しているため，2 乗誤差基準は誤り率 ($0 \leq r \leq 1$) となり，0 の時が全て正解，1 の時は全て誤りを意味する．

次に，求めた \hat{S}^T の時刻毎の平均

$$E(\hat{s}_t) = \frac{1}{T} \sum_{t=1}^T \hat{s}_t \quad (2)$$

は聴取者のこの楽曲に対する嗜好を表す値と考えられるため，これを楽曲の推薦値 ($0 \leq E(\hat{s}_t) \leq 1$) と呼ぶことにする．

ここで適切に推薦値のしきい値を定めれば，そのしきい値よりも大きい推薦値の楽曲は聴取者に推薦可能であると判断できる．そこで本研究では次のようにしてしきい値を定めることにした．

1. 外れ値を取り除くためにサンプリングした楽曲の中から 2 乗誤差基準 (1) が 0.5 以下の楽曲だけを取り出す．
2. 取り出した N 個の楽曲それぞれに対して推薦値 (2) を求める．
3. 各楽曲の推薦値を昇順でソートし，順に $X_i (i = 1, 2, \dots, N)$ とする．また $X_0 = 0$ および $X_{N+1} = 1$ とする．
4. 範囲 D_i を $D_i = \{x : X_i \leq x < X_{i+1}\}, (i = 0, 1, 2, \dots, N)$ と定義する．
5. しきい値をある値に固定したときに，推薦可能と判断された楽曲を聴取者が実際に好きである割合を取り出した N 個の楽曲から求めてこれを正解率と呼ぶことにする．このとき D_i の範囲内では正解率は一定となる (図 2) ．
6. 範囲 D_i の中で正解率が最大，かつ，その範囲が最大であるものを求める．例えば図 2 では範囲 D_3 と D_6 の正解率が最大となるが，より範囲が広い D_6 の方を選択する．
7. 求めた範囲の中心の値を推薦値のしきい値とする．

こうして求めたしきい値を用いれば，印象変化をサンプリングしていない楽曲でも聴取者に推薦できるか判断することが可能になる．

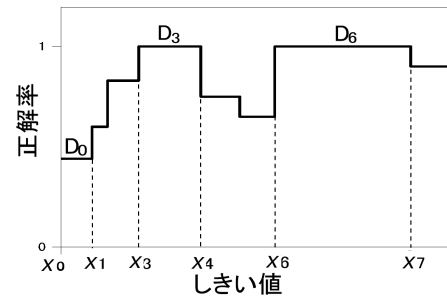


図 2: しきい値と正解率の関係例

5 まとめ

本研究では聴取者の楽曲に対する印象変化をもとにして楽曲を推薦するシステムを構築した．このシステムは，楽曲のコード進行と聴取者の印象変化に HMM モデルを適用し，対象となる楽曲の推薦値を求め，その値が適切に定めたしきい値よりも大きい場合には推薦可能として聴取者に提示するものであった．

今後は，実際に実験をおこなって本推薦システムの性能評価をおこなう予定である．また他の推薦システムとの比較もおこなう予定である．

参考文献

- [1] 帆足啓一郎, 上月勝博, 菅谷史昭, 楽曲配信サービスを支える音楽情報検索技術, 電子情報通信学会誌, pp.529-534, Vol.88, No.7, 2005.
- [2] 千田一考, 藤澤公也, 印象の類似度に基づくオンライン楽曲推薦手法の研究, 情報処理学会第 69 回全国大会講演論文集 (2), pp.187-188, 2007.
- [3] Cohen.W.W, Fan.W, Web-collaborative filtering: recommending music by crawling the Web, Computer Networks, pp.685-698, Vol.33, 2000.
- [4] 石先広海, 帆足啓一郎, 菅谷史昭, 甲藤二郎, ユーザ嗜好に基づく音楽情報検索システムにおける学習データ抽出手法, 情報処理学会研究報告, pp.73-78, 2006-MUS-64, 2006.
- [5] 辻康博, 星守, 大森匡, 曲の局所パターン特徴量を用いた類似曲検索・感性語による検索, 電子情報通信学会技術報告, pp.17-24, SP96-124, 1997.
- [6] 東海林智也, HMM によるリスナーの印象変化推定と楽曲推薦システムへの応用, 情報処理学会第 69 回全国大会講演論文集 (2), pp.121-122, 2007.
- [7] 北研二, 確率的言語モデル, 東京大学出版会, 1999.