

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

Estimation of Change of Listener's Impression Using HMM and Its Application to Music Recommendation System

東海林智也*

函館工業高等専門学校

Tomoya Tokairin

Hakodate National College of Technology

1 はじめに

インターネットや携帯電話が普及して楽曲のダウンロード販売が一般的になるにつれ、楽曲のジャンルやアーティスト名など各楽曲が持つ固定的な情報だけに頼らずに、各リスナーの楽曲に対する個人的な印象(嗜好)にもとづいて楽曲を検索する楽曲推薦システムの構築が盛んにおこなわれている [1].

従来では、各楽曲に対する印象が似ている他人の購買履歴を利用して楽曲の検索と推薦をおこなう協調フィルタリング手法 [2], 楽曲のジャンル情報や信号に含まれている MFCC などの特徴量とリスナーの楽曲に対する印象情報を、正準相関分析やクラスタリングなどの統計手法を用いて結びつけて検索と推薦をおこなう手法 [3][4] などが提案されているが、いずれの場合も楽曲全体を通したリスナーの印象をもとにして楽曲の検索と推薦をおこなっているものが多い.

しかしリスナーの楽曲に対する印象は楽曲の進行によって変動するため、リスナーが常に同じ印象を持ち続けることは一般的には無い. 例えば、ある楽曲のイントロパートは嫌いでもサビパートは好きであるといった状況は良く見られる. 従ってリスナー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によってより効果的な楽曲の検索と推薦が可能になると期待できる.

そこで本稿では、楽曲の進行につれて変化するリスナーの楽曲に対する印象変化を時系列データとみなして HMM (Hidden Markov Model: 隠れマルコフモデル) により推定し、その推定した印象変化をもとにして楽曲推薦システムを構築することを提案する. さらに実験をおこなって提案手法の有効性を検証する. なお、本稿では楽曲の進行情報としてコード進行を使用する.

2 コード進行と印象変化のモデル化

2.1 コード進行に関する時系列データ

はじめに楽曲の進行情報、すなわちコード進行に関する情報を次のようにして形式化する. 一般的には連続する 2 つのコードをひとつの組としてコード進行を考えることが多いため、楽曲に含まれるコードを先頭から 2 つずつ分割して取り出し、それらの集合を時系列データと考えて楽曲のコード進行データとする. 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データとなる.

次にこのコード進行データを数値化する. メジャーコード、マイナーコード、セブンスコード、転回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する必要がある. そこでトニックの根音と同じ根音を持つコード全てに状態番号 0 を割り当て、トニックの根音から k だけ半音上の音を根音とするコード全てに状態番号 k を割り当てることにすると、全てのコードに対して 0 から 11 までの状態番号が割り振られる. 例えば上で挙げたコード進行の場合は、C メジャースケールのもとでは $\{\{0, 5\}, \{7, 0\}, \{5, 0\}\}$ となる.

最後に、時刻 t , ($1 \leq t \leq T$) におけるコード進行を c_t とおいてコード進行データを $C^T = c_1 c_2 \cdots c_T$ という時系列データの形式で表すことにする. 例えば上で挙げたコード進行の場合は $C^3 = c_1 c_2 c_3$, $c_1 = \{0, 5\}$, $c_2 = \{7, 0\}$, $c_3 = \{5, 0\}$ となる.

2.2 印象変化に関する時系列データ

リスナーの好き、嫌いなどの楽曲に対する印象変化情報に関しては次のようにして形式化する. 状態 (印象) 数として M 個を選んだとき、各状態に 0 から $M-1$ 番までの数値を割り当てる. 今回は問題を単純にするために「好き」(状態番号 1)、「普通」(状態番号 0) の $M=2$ 状態とする.

ここでコード進行データと同様に、時刻 t , ($1 \leq t \leq T$) における印象を s_t とおいて印象変化データを $S^T = s_1 s_2 \cdots s_T$ という時系列データの形式で表すことにする. 例えば、リスナーがある楽曲を聴いたときに、先頭から好き 普通 普通 好きという印象変化が現れたとすると、その印象変化データは $S^4 = s_1 s_2 s_3 s_4$, $s_1 = 1, s_2 = 0, s_3 = 0, s_4 = 1$ となる.

2.3 コード進行と印象変化の同時確率

実際にリスナーの印象変化をサンプリングをするときは、コード進行データ C^T の状態 c_t が切り替わる時点とサンプリングのタイミングを一致させる. したがって、ひとつの楽曲からコード進行データと印象変化データの時系列の集合 $\{C^T, S^T\}$ が得られる.

ここで、 $\{C^T, S^T\}$ の同時確率を $P(C^T, S^T)$ とする. S^T が未知、かつ C^T が既知であるときに、この同時確

*tokai@hakodate-ct.ac.jp

率を最大にする \hat{S}^T が S^T の最尤推定量となる。

2.4 HMMによるモデル化

$P(C^T, S^T)$ を最大化する \hat{S}^T を求めるために HMM を適用する。本稿では 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 の様な HMM を構成した。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。これはリスナーを楽曲の作曲者と見立てたときに作曲者の観測不能な心理的印象変化に従ってコードが出力されたというモデルを考えていることに相当する。なお, 時刻 $t+1$ におけるコード進行 c_{t+1} の出現確率は, 実際には s_{t+1} だけではなく c_t の影響も受けるが, 条件数が増大することから今回は c_t と c_{t+1} は独立であるとした。

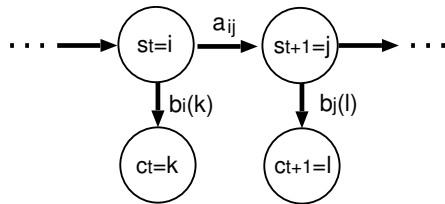


図 1: 提案する HMM

3 実験による検証

3.1 実験概要

リスナーがいくつかの楽曲を聴いたときの印象変化をサンプリングし, 印象変化の遷移確率 a_{ij} と, ある状態 (印象) i におけるシンボル (コード進行) k の出現確率 $b_i(k)$ を推定する。次に確率の推定に用いなかった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 を図 1 の HMM をもとに求めて実際の S^T と比較する。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。今回は状態数として好き (1), 普通 (0) の 2 通りのみを使用しているので 2 乗誤差基準は誤り率 ($0 \leq r \leq 1$) を表している。

なお今回の実験では楽曲としてメジャースケールの童謡 25 曲を用いる。また HMM の内部状態の推定アルゴリズムとして viterbi アルゴリズム [5] を使用する。

3.2 結果

実験に用いた楽曲のうち 1 曲を取り除き, 他の楽曲を用いて遷移確率やコードの出現確率を推定し, 取り除いた楽曲の \hat{S}^T を求めて 2 乗誤差基準を計算した。これを各楽曲に対して繰り返し, 2 乗誤差基準の平均と標

準偏差を求めた。結果は平均が 0.33, 分散が 0.16 であり, 印象変化がうまく推定されていることが分かった。

4 楽曲推薦システムへの応用

提案手法の楽曲推薦システムへの応用を考える。ある楽曲から求めた印象変化の推定値 \hat{S}^T の時刻毎の平均

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

を考えると, あるしきい値よりも大きい値の楽曲は好きであるコード進行が多いということを意味している。そこで式 (2) が大きい楽曲から順にユーザに提示することによって楽曲推薦システムを構築できる。

5 問題点

提案手法の問題点として, コードが進行するたびに楽曲の再生を止めてリスナーの印象を調べる必要があるためサンプリング時間が長いことが挙げられる。解決法としては既存の楽曲ではなく, 人工的に作成した楽曲を用いて効率的にコード進行と印象変化の情報を収集することなどが考えられる。

6 まとめ

本稿では楽曲の進行につれて変化するリスナーの楽曲に対する印象変化を HMM を用いて推定する手法を提案し, 実験によりその有効性を実証した。また, 推定した印象変化をもとにして楽曲推薦システムを構築することを提案した。

今後の課題としては従来の楽曲推薦システムと今回提案した推薦システムの性能比較や, 問題点として挙げたサンプリング時間の短縮などがある。

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] Cohen.W.W, Fan.W, Web-collaborative filtering: recommending music by crawling the Web, Computer Networks, pp.685-698, Vol.33, 2000.
- [3] 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.
- [4] 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.
- [5] 北研二, 確率的言語モデル, 東京大学出版会, 1999.