

WEB コミュニケーション連動楽曲再生システム

石先 広海†

帆足 啓一郎†

小野 智弘†

†KDDI 研究所

1 はじめに

現在，WEB を介して様々なマルチメディアコンテンツサービスを利用することが可能となり，ユーザはPCや携帯端末などを利用してコンテンツを視聴しながらWEB上でコメントをやりとりするなどのWEBコミュニケーションが活発化している．例えば，ビデオ共有サイトやインターネットラジオでは，コンテンツを視聴しながら，コンテンツに対する感想や意見をコメントとして投稿し合うことでコミュニケーションをとることが可能である．

一方で，上記のサービスでは主にコンテンツに対するコメントを投稿することが一般的で，WEBコミュニケーションに適した楽曲をBGMとして再生する環境は存在していない．例えばSNSでは，共通のトピックに対して興味を持つユーザ同士が意見交換や情報共有のためにコメントを投稿をすることがある．これらのトピックやコメントの内容に適した楽曲をBGMとして再生することで，コメントやコミュニケーション場全体の情景表現力を向上させた新しい音楽聴取環境の実現が期待できる．

そこで本稿では，音楽とテキストを連動させた新しい音楽聴取環境を実現するために，WEB上で音楽を聴取しながらコメントをやりとりし，コメントと連動して音楽が再生されるなど，WEBコミュニケーションに適した楽曲を自動で再生するシステムを提案する．さらに主観評価実験により本システムの有効性を検証する．

2 提案システム

本システムでは，コメントに含まれる単語と，楽曲の印象および感情情報のマッチングにより楽曲を選定することで，WEB上でのコミュニケーションに適した楽曲をBGMとして自動で再生する．提案システムの概要図を図1に示す．

楽曲特徴量生成部では，楽曲の歌詞および音響データを利用して楽曲特徴量を生成し，楽曲DBに格納する．楽曲選定処理部では，入力したWEBコミュニケーションデータから，形態素解析により文書特徴量を生成し，楽曲特徴量との関連度に基づいて楽曲を選定し，

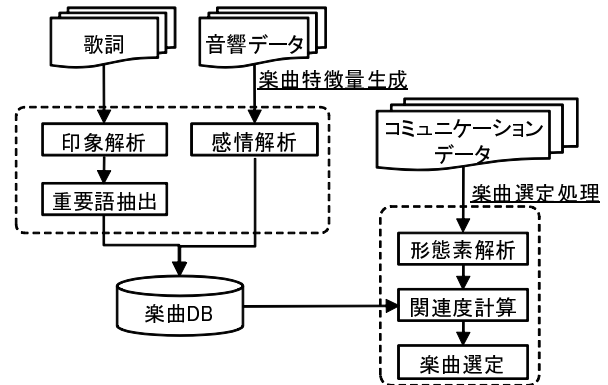


図1: 提案システム概要図

BGMとして再生する．楽曲の印象および感情情報を利用することで，トピック内の感情及び印象語に関連した楽曲をBGMとして再生でき，コメントやコミュニケーション場の情景表現力の向上が期待できる．以下に楽曲特徴量生成および楽曲選定処理の詳細を記述する．

2.1 楽曲特徴量生成処理

本システムでは，4つのカテゴリ（歌詞に対する印象，音響的特徴に基づく感情情報，タイトル，歌詞重要語）から楽曲特徴量を生成する．まず，歌詞に対する印象情報は，文献[1]に記載の楽曲分類方法を利用する．文献[1]では，楽曲歌詞をTF * IDF法によりベクトル表現に変換し，学習データとなる歌詞印象情報を基にしてSVMにより印象情報を付与している．本システムでは全ての楽曲に対して，“季節”，“時間帯”，“天候”情報を推定し，印象情報として付与する．

音響的特徴に基づく感情情報は，文献[2]に記載されている音響的特徴からの感情推定方法を利用する．文献[2]では，感情ラベルが付与された学習データから感情空間を構築し，感情空間に対する分布に基づいて，5つの感情情報（“怒り”，“怖い”，“嬉しい”，“悲しい”，“優しい”）を付与できる．さらに追加の特徴量として，楽曲のタイトル情報とタイトルに含まれる名詞及び形容詞，歌詞からTF * IDF法により抽出した重要語を楽曲の特徴量として利用する．

Music Playing System Associated with WEB Communication.

†Hiromi ISHIZAKI †Keiichiro HOASHI †Chihiro ONO

†KDDI R&D Laboratories Inc.

2.2 楽曲選定処理

入力データに対する形態素解析により、名詞および形容詞情報を抽出し、文書特徴量 ($W = \{w_1, w_2, \dots, w_K\}$) とする。生成した文書特徴量と楽曲特徴量の関連度は以下の式により算出する。

$$score = \sum_{j=1}^{N_{Ct}} \frac{\alpha_j \sum_{i=1}^K count(Ct_j, w_i)}{N_j} \quad (1)$$

ここで、 N_{Ct} はカテゴリ数、 $count$ はカテゴリ内の単語 Ct_j と W の一致個数を計算する関数とし、 α は各カテゴリの重み係数、 K は文書特徴量の要素数、 N は各カテゴリの要素数とする。最終的に最もスコアの高い楽曲を BGM として選定し、再生する。

3 主観評価実験

第2章に記載の印象および感情情報を利用して楽曲を選定する方式の有効性を検証するため、被験者による主観評価実験を実施した。本実験では、式(1)における重み係数 α を調整し、印象及び感情情報(手法1)タイトル情報(手法2)、歌詞重要語(手法3)、の各要素によって楽曲を選定する3つの手法を用いて比較を行った。

3.1 実験方法

被験者に対してコミュニケーションデータを提示し、被験者がデータを閲覧する際に各手法によって選定された3種類の楽曲を BGM として再生した。被験者(7名)は、聴取した全ての BGM に対して5段階評価(5:BGM としてとても適している, 1:BGM としてまったく適していない)を基準として評価値を付与した。

提示したデータは、SNS の5ジャンル(音楽、映画、スポーツ、旅行、地域)から各2項目のトピックを選出し、トピックごとにトピック作成時の投稿コメントおよびそれに対するコメント15件(投稿順)を抜粋して作成した(合計10データ)。評価実験に用いた楽曲は市販 CD から収集した J-POP (3,407 曲) を利用し、全ての楽曲に対して2.1節に記載の方法で楽曲特徴量を作成した。形態素解析器は Mecab¹ を用いた。

3.2 実験結果

図2に、主観評価値の全体平均および各コミュニケーションデータの評価値平均を示す。各手法の全体平均値は、それぞれ3.2, 2.7, 2.9 となり、印象及び感情情報を用いた手法1の平均値が最も高い結果となった。ま

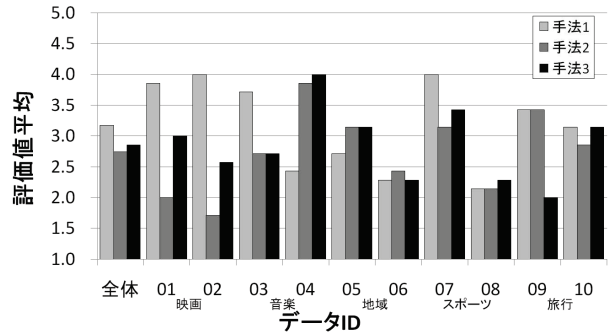


図2: 主観評価実験結果: 全体平均および各ジャンルの評価値平均

た、標準偏差はそれぞれ 0.95, 1.03, 1.07 となり、手法1が最も安定していることが分かる。

手法1では、コミュニケーションデータのトピックが「映画の好きな台詞」「スキー場情報」など、印象及び感情情報が含まれやすいテーマであった場合に評価値が高くなる傾向が見られた。実験結果よりコミュニケーションに適した BGM を選定する場合に、印象および感情情報が有効であることが示唆された。

一方で、「自身の思い出に関する楽曲」「地域のおすすめ情報」などのように、被験者に明確な楽曲のイメージを想起させる場合や、印象・感情語が含まれない場合には評価値が低い傾向があった。提案システムの重みパラメータや楽曲特徴量の形式を改善し、タイトル情報や歌詞重要語の特徴を考慮することで改善が可能であると考えられる。

4 まとめ

本稿では、楽曲の歌詞と音響的特徴から推測した印象・感情情報を利用して楽曲を選定する WEB コミュニケーション連動楽曲再生システムを提案した。評価実験により、WEB コミュニケーションの BGM 選定する際に印象・感情情報が有効であることが示唆された。今後は、複数の指標を統合させて効果的な楽曲選定が可能なシステムを検討していく。

参考文献

- [1] 舟澤慎太郎, 他:歌詞の印象に基づく楽曲検索のための楽曲自動分類に関する検討, 第71回情報処理学会大会, 5R-2, 2009.
- [2] Eerola, T., et.al: Prediction of multidimensional emotional ratings in music from audio using multivariate regression models., ISMIR2009.

¹Mecab: <http://mecab.sourceforge.net/>