

コミュニケーションメディアとしてのプレイリストを目指して For Realization of Playlist-Mediated Communication

梶 克彦*

平田 圭二†

長尾 確‡

1. はじめに

近年大容量の携帯音楽プレイヤーが登場し、インターネット上に多数の音楽コンテンツが存在するようになり、いつでもどこでも容易に音楽を楽しむことが可能になった。またネットラジオなどのデジタルコンテンツを自動的に iPod に転送する Podcasting というサービスが現在利用可能である。自分に合ったコンテンツを、インターネット上の膨大なコンテンツの中から自動的に携帯音楽プレイヤーに取り込みたいという要求が高まっている。

そのため楽曲推薦やプレイリストの自動生成に関する研究が盛んに行われている。既に iTunes Music Store などのインターネットを通じた音楽配信サービスがいくつか利用可能である。また近い将来、定額制サービスが音楽配信にも適用されるだろう。そのような環境下では、多くのユーザ間において、楽曲の ID のみで構成されるプレイリストが共有されることになるだろう。

また現在 Blog や SNS(Social Networking Service) により、インターネット上でのテキストや絵、写真などによる活発な個人間コミュニケーションが行われている。音楽は多くの人にとって容易に制作できないメディアであるが、プレイリストはテキストや写真などと同様、容易に制作することができ、また制作者の意図や感情を込めることができる。我々は今後プレイリストがテキスト・絵・写真と同様、重要なユーザ間のコミュニケーションメディアになると考えている。以降プレイリストを介したコミュニケーションを Playlist-Mediated Communication と呼ぶことにする。

従来の楽曲推薦に関する研究は、協調フィルタリングによるプレイリスト生成システム [1][2] やジャンル、アーティストなどの情報を利用したプレイリスト生成 [3] などがある。しかし、楽曲推薦に有効である情報の中にはいくつかの収集することが困難である主観的・個人的な情報が存在する [8]。そのためそういった情報を直接的に付与する手法が必要となる。そこで我々はアノテーションの技術に着目した。セマンティック Web の研究分野では、アノテーションは柔軟にコンテンツに対して個人的・公共的なコメントなどのメタデータを記述する手法とし

て広く認知されている [4][5]。そのためアノテーションは主観的・個人的な情報を柔軟に収集するのにふさわしい技術であるといえる。

Playlist-Mediated Communication を実現するために必要な要素は、プレイリストの制作・推薦・鑑賞といった活動を活発にすることであると考え、その実現の第一歩として、プレイリスト推薦の手法を提案する。本システムの実現により、プレイリスト制作・推薦の促進を実現することができるであろう。プレイリストをコミュニケーションメディアとして利用できるようにするためには、あるユーザが制作したプレイリストを、他のユーザが容易に参照できる必要がある。そこで本システムを Web サービスとして設計する。Web サービスであれば、インターネットを介した他のユーザとの円滑な意見交換やプレイリスト推薦の実現が期待できる。

本論文は以下のように構成される。まず 2 章でプレイリスト推薦システムの詳細について述べ、3 章では構築したプレイリスト推薦システムの評価実験の結果を述べる。4 章では本論文をまとめ、今後の課題と展望について言及する。

2. プレイリスト推薦システム

ジャンルやアーティスト、歌詞など、一般的に楽曲推薦に利用すると効果的であるとされている楽曲の特徴量がいくつか存在する [6][7]。一方、日常生活における音楽との接し方から推測されるように、リスナの置かれている状況が、推薦すべき楽曲の選択に強く依存するのではないかと考えられる。そこで我々はリスナの嗜好と状況に合ったプレイリスト推薦システムを構築する。本研究では、楽曲の特徴量として歌詞・表現している情景(楽曲情景)・聴きたい状況(鑑賞状況)の 3 種類を採用した。

2.1. 歌詞とアノテーションによる楽曲間類似度推定

楽曲情景と鑑賞状況については、鑑賞するリスナの解釈に強く関わる情報であるため、楽曲の音響情報を自動解析して得ることが困難である。そこで、別途用意した複数項目のアンケート形式のアノテーションシステムにより、各楽曲に対して楽曲情景と鑑賞状況の情報を収集し、その結果を楽曲の特徴量として利用した。

*名古屋大学大学院 情報科学研究科

†NTT コミュニケーション科学基礎研究所

‡名古屋大学 情報メディア教育センター

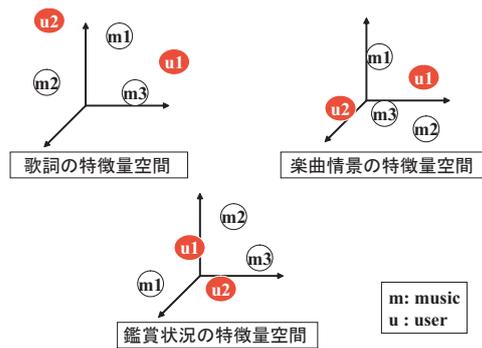


図 1: 歌詞・楽曲情景・鑑賞状況の特徴量空間

収集するアノテーションは、リスナの主観に依存するあいまいな情報であるため、多くのユーザからそれらの情報を取得し、統計的に処理を行う必要がある。そこで別途アノテーションシステムを構築し、各楽曲のアノテーションを収集した。一つの楽曲に複数のユーザによるアノテーションが付与されている場合はそれらのアノテーションの平均をその楽曲の特徴量として利用する。同時に歌詞を利用し、各楽曲の TF*IDF[10] を計算しておく。 l_{m_i} 、 c_{m_i} 、 s_{m_i} を歌詞・楽曲情景・鑑賞状況の各特徴量空間における楽曲 m_i の特徴量とすると、 l_{m_i} は以下に示す式によって求められる。

$$l_{m_i} = (f(m_i, k_0), f(m_i, k_1), \dots, f(m_i, k_n))$$

TF*IDF によって求められたキーワードの数は n であり、 k_i は i 番目のキーワードを表している。 $f(m_i, k_j)$ は楽曲 m_i の歌詞が持つキーワード集合中に楽曲 k_j が持つキーワード集合の中のキーワードが多く含まれている場合に 1 に近づき、全く含まれていないと 0 になる。

図 1 は、各楽曲が持つ歌詞・楽曲情景・鑑賞状況の特徴量から、各特徴量空間にマップされていることを表している。各楽曲は m_1 、 m_2 、 m_3 として表されており、 u_1 、 u_2 はこれらの特徴量空間にマップされたユーザを表している。各特徴量空間へのユーザのマッピングについては、次節について詳細に述べる。以上の処理により、各楽曲間のコサイン距離を計算することができる。コサイン距離は、ある特徴量空間における 2 要素間の類似度を計算するために有効であり、値が大きいほど類似しているといえる。楽曲 m_i と m_j の類似度 $\text{sim}(m_i, m_j)$ は以下の式で算出される。

$$\text{sim}(m_i, m_j) = \alpha \cos(l_{m_i}, l_{m_j}) + \beta \cos(c_{m_i}, c_{m_j}) + \gamma \cos(s_{m_i}, s_{m_j})$$

α 、 β 、 γ は歌詞・楽曲情景・鑑賞状況の特徴量空間における重みであり、 $\cos(a, b)$ は a 、 b 間のコサイン距離を表す関数である。複数の特徴量のうちどの特徴量を強調するかによって α 、 β 、 γ の値は変化させることができる。今回のプレイリスト生成システムでは、ユーザの置かれた状況に、より適した楽曲を選出することができるよう

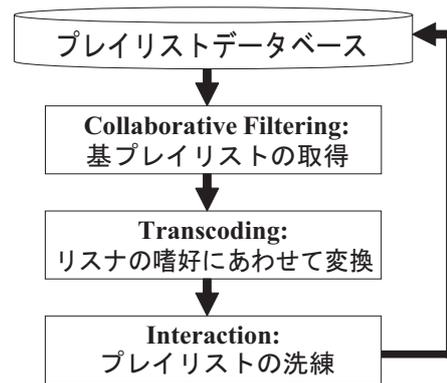


図 2: プレイリスト生成の流れ

にするため、 α 、 β に比べ γ の値を大きく設定した。他の楽曲推薦に有効であるとされているジャンルやアーティストなどのメタデータに関して、特徴量空間を追加することにより本システムに導入することが可能である。

2.2. プレイリスト生成の流れ

図 2 はプレイリスト生成の流れを表している。まず多くのプレイリストの中からユーザに適したプレイリストを選出することができる協調フィルタリングを行う。次に選び出されたプレイリストを、よりリスナの嗜好に適合させるトランスコーディングを施しユーザにプレイリストを提示する。さらにシステムとリスナのインタラクションによりリスナの嗜好に適応していく。以下に詳細を述べる。

システムはそれぞれのユーザのプレイリスト生成履歴から、そのユーザの嗜好に合った楽曲集合を保持している。ここで、ユーザにとって最も嗜好に合う楽曲が存在し、その楽曲はユーザの嗜好に合った楽曲集合を元に推測できると仮定した。そして歌詞と楽曲情景の特徴量に関してそれぞれユーザの嗜好に合った楽曲集合の平均を計算し、図 1 のようにそれぞれの特徴量空間にマップする。また、ユーザの置かれている状況に関しては、本システムの鑑賞状況入力フォームから入力を受け付け、その際にユーザを鑑賞状況の特徴量空間にマップする。鑑賞状況入力フォームの項目は、2.1 節で述べた各楽曲に対するアノテーションシステムにおける鑑賞状況の項目と同等である。

プレイリスト生成を行う際に、まずプレイリストデータベースから協調フィルタリングにより基プレイリストを選び出す。類似するユーザを発見するために、ユーザ間の類似度を歌詞と楽曲情景の特徴量空間から算出し、さらに類似するユーザが生成したプレイリストから、ユーザ間の類似度、鑑賞状況の類似度を踏まえて一つの基プレイリストを選出する。

次に基プレイリストをさらにユーザの嗜好に適合させるトランスコーディングを行う。具体的には、ユーザが過去に嗜好に合わないとシステムにフィードバックした

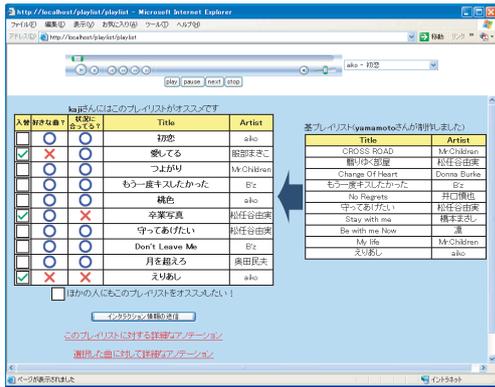


図 3: 提示されるプレイリストの例

楽曲を取り除き、楽曲データベースから替わりの楽曲を選んで追加する。同時に、ユーザが鑑賞経験のない楽曲を3割の割合で含ませている。聴き慣れた楽曲の中に適度に鑑賞経験のない楽曲を挿入することで、ユーザの嗜好の幅を徐々に広げる効果が期待できる。嗜好に合わない楽曲などの替わりに追加する楽曲は、ユーザの嗜好と、置かれた状況から、各特徴量空間におけるユーザと楽曲の類似度を計算し、選出される。ユーザと楽曲間の類似度は、2.1節で述べた楽曲間の類似度計算と同様の計算式から導き出される。

以上の処理を経てユーザには図3のようにプレイリストが提示される。協調フィルタリングにより選出された基プレイリストが右側に、トランスコーディング後のプレイリストが左側に表示されている。ユーザはプレイリスト上部に埋め込まれたプレイヤから、通常の音楽プレイヤーと同様の操作でプレイリストを楽しむことができる。

ユーザは実際にプレイリストを鑑賞し、各楽曲が嗜好に合っているか、状況に合っているかといった情報をフィードバックすることができる。システムはユーザからのフィードバックを受け取り、気に入らない楽曲の入れ替えを行い再度ユーザにプレイリストを提示する。同時にユーザのプロファイルを更新する。具体的には、好きな曲かどうか、状況に合っているかどうかといった情報から、ユーザの特徴量空間の基底ベクトルの変換を行う。以下は歌詞の特徴量空間における基底ベクトル変換の式である。

$$\text{base}_u = \text{normalize}(\text{base}_u + \delta \frac{l_f - l_u}{|l_f - l_u|} - \delta \frac{l_d - l_u}{|l_d - l_u|})$$

base_u はユーザ u の歌詞の特徴量空間における基底ベクトルである。 l_f は嗜好に合っているとフィードバックされた楽曲の歌詞特徴量の平均であり、 l_d は嗜好に合わないという楽曲の平均である。基底ベクトルの変換レートは δ で表されている。基底ベクトル変換により、特徴量空間を嗜好に合った楽曲の方向に短縮し、嗜好に合わない楽曲の方向に拡張している。楽曲情景と鑑賞状況の特徴量空間の基底ベクトルについても、同様に変換を行う。

さらに、ユーザが今までに鑑賞した曲や、好きな曲、嫌いな曲、どのような状況でどのようなプレイリストを生成したかといった情報を保持し、各特徴量空間へのユーザのマッピング、協調フィルタリング、トランスコーディングの際にそれらの情報を利用している。

2.3. 鑑賞履歴の取得とユーザプロファイルへの反映

コンテンツ推薦を行う際、一般にユーザプロファイルに導入すべき情報として、どのコンテンツを、どれだけ利用したかという情報が挙げられる。本システムではプレイリストプレイヤーが埋め込まれており、提示したプレイリストをどのように聴いたかという鑑賞履歴を取得することができる。プレイリストプレイヤーはスタート・ストップ・楽曲選択などの操作情報を随時、システムに通知し、ユーザがどの楽曲のどの部分を何秒間鑑賞したかという詳細な情報をプロファイルに反映させている。

このような鑑賞履歴の取得は、ユーザプロファイルへの反映だけでなく、楽曲へのアノテーションとして利用することができる。ユーザの楽曲解釈のアノテーションはそれぞれの楽曲ごとにその情報量が異なることが多い。そこで、アノテーションが十分に集まっていない楽曲について、鑑賞履歴を利用し、情報量を補うことが可能である。例えば、似たような状況において多くの人が繰り返し聴いている楽曲であれば、その状況に合った楽曲といえることができるであろう。

3. 評価実験

評価実験に利用した楽曲はRWC音楽データベース [9] のPOPS100曲と、一般のPOPS130曲である。事前に各楽曲に対して、2.1節で述べた楽曲情景・鑑賞状況に関するアノテーションを収集し、その情報を楽曲の特徴量として利用した。

構築したシステムの評価を行うため、7人の被験者に対して以下の手順で評価実験を行った。まず自由に自分の置かれている状況を設定してもらい、本システムの鑑賞状況入力フォームから入力してもらおう。システムは各被験者の嗜好と、入力した鑑賞状況に適したプレイリストを提示する。そこで実際にプレイリストを鑑賞してもらい、各楽曲に対して嗜好に合っていたか、状況に合っていたかの情報をチェックしてもらい、システムにフィードバックしてもらおう。システムは受け取った情報を基にプレイリスト作り変え、再びユーザにプレイリストを提示する。以上の行為をプレイリストのすべての楽曲が嗜好・状況共に合ったものになるまで繰り返してもらおう。以上のプレイリスト生成プロセスを、3回繰り返してもらった。

本システムがユーザの嗜好・状況に合っているプレイリストを生成しているかどうかを確認するため、各プレイリスト生成プロセスにおいて、鑑賞状況入力後提示されるプレイリスト中に、嗜好に合う楽曲・状況に合う楽

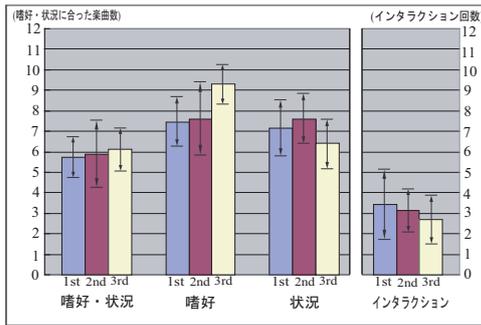


図 4: 評価実験の結果

曲・嗜好と状況に合う楽曲はそれぞれ何曲存在したか、またすべての楽曲が嗜好と状況に合うまで行ったインタラクションの回数を評価した。評価実験の結果を図 4 に示す。それぞれのグループは 3 つのバーによって構成されており、左から 1 回目、2 回目、3 回目のプレイリスト生成プロセスにおける評価値である。各バーの頂点の矢印は各評価値の標準偏差を表している。

図 4 の左から 2 番目に表される嗜好に合った楽曲数のグラフと、左端の嗜好・状況共に合った楽曲数のグラフからは、プレイリスト生成プロセスを重ねるにつれてユーザに合った楽曲数が増加していることが読み取れる。また右端のインタラクションの回数は、徐々に減少していることがわかる。このことから本システムを長く利用するユーザほど、嗜好と状況に合ったプレイリストを、少ないインタラクションにより得ることができるといえる。

一方、左から 3 番目のグループに現れている、状況に合った楽曲数のグラフでは、1 回目、2 回目、3 回目とプレイリスト生成を繰り返すことによる増加傾向は見られない。鑑賞状況に合った楽曲提示については、プレイリスト生成を繰り返すことにより改善されているとはいえない結果であった。

この結果は以下の 2 つの問題点が原因となって起こったと考えられる。一つ目は本システムで入力を要求する鑑賞状況項目の組み合わせが、利用できる楽曲数に比べて非常に多いという点、二つ目に、各楽曲に対して収集した楽曲情景・鑑賞状況のアノテーションは各楽曲に均一に付与されているわけではないという点である。

前者の問題点に対しては、より多くの評価実験を繰り返し、鑑賞状況項目をより効果的な項目に絞り込み、同時に利用できる楽曲数を追加することで対応できると考えられる。後者の、情報の偏りに関する問題は、一般ユーザを対象としたアノテーションシステムにおいては、コンテンツの人気などにより頻繁に発生する問題である。この問題を解決するために、ユーザをアノテーションが不足しているコンテンツに誘導したり、2.3 節で述べたようにリスナの鑑賞履歴などを利用することでアノテーションの均一化を図る仕組みが必要となる。

しかし鑑賞履歴のような情報は、アノテーションシス

テムから付与された情報と比べて一般的に信頼性に劣るため、そういった情報を適用するためには、各アノテーションの重みを考慮する必要がある。

4. 終わりに

プレイリストにまつわる制作・推薦・鑑賞といった活動を活発にすることで Playlist-Mediated Communication を実現することができると考え、その第一歩として、プレイリスト推薦システムを構築した。また本システムの評価実験を通して、アノテーションがユーザの心的状態や置かれている状況、またユーザと楽曲間の関係を正確かつ動的に表すために有効であると確認された。同時に、協調フィルタリング・トランスコーディング・インタラクションの組み合わせがユーザの嗜好に合ったプレイリスト推薦に有効であると確認できた。本システムにより、プレイリスト制作と推薦を促す仕組みが実現できたと考えるだろう。

今後の課題を以下に述べる。プレイリストをコミュニケーションメディアとして利用するためには、プレイリストに対するコメントや、そのプレイリストの他のコンテンツとの関連など様々な付加情報の存在が重要となる。そこで我々はプレイリストに対してそういった情報を付与できるアノテーションシステムを構築中である。本アノテーションシステムを通して、ユーザはプレイリストに対する意見など様々な情報を記述することができ、また他のユーザのコメントを閲覧することができる。プレイリストに付与されたアノテーションを利用することにより、プレイリスト生成手法も改良することができるだろう。また、プレイリスト制作・推薦のさらなる促進手法に加え、積極的なプレイリスト鑑賞を促す手法について検討していく予定である。

参考文献

- [1] M. Anderson, M. Ball, H. Boley, S. Greene, N. Howse, D. Lemire and S. McGrath, "RACOFL: A Rule-Appling Collaborative Filtering System," Proceedings of COLA, 2003.
- [2] U. Shardanand and P. Maes, "Social Information Filtering: Algorithms for Automating "Word of Mouth",," Proceedings of CHI, 1995.
- [3] S. Pauws and B. Eggen, "PATS: Realization and user evaluation of an automatic playlist generator," Proceedings of ISMIR, 2002.
- [4] Katashi Nagao, "Digital content annotation and transcoding," Artech House Publishers, London, 2003.
- [5] J. R. Smith and B. Lugeon, "A Visual Annotation Tool for Multimedia Content Description," Proceedings of the SPIE Photonics East, Internet Multimedia Management Systems, pp.49-59, 2000.
- [6] E. Daniel, B. Whitman, A. Berenzweig and S. Lawrence, "The Quest for Ground Truth in Musical Artist Similarity," Proceedings. of ISMIR, 2002.
- [7] B. Logan, A. Kositsky and P. Moreno, "Semantic Analysis of Song Lyrics," IEEE International Conference on Multimedia and Expo (ICME), 2004.
- [8] J. H. Lee and J. S. Downie, "Survey of Music Information Needs, Uses, and Seeking Behaviours: Preliminary Findings," Proceedings of ISMIR, 2004.
- [9] Masataka Goto, "Development of the RWC Music Database," Proceedings of the 18th International Congress on Acoustics (ICA 2004), pp.1-553-556, 2004.
- [10] G. Salton and C.S. Yang, "On the specification of term values in automatic indexing," Journal of Documentation, pp.29:351-372, 1973.