

ユーザの好みを考慮する端末適応型音楽エージェントシステムの提案

竹澤 直人[†] 勝本 道哲^{††}

[†]電気通信大学 ^{††}郵政省通信総合研究所

アブストラクト

近年インターネットを通じて楽曲が配信されたり CD が販売されたりしており、将来ほとんどの楽曲がインターネットを介して CD 以上の品質で配信されると予想できる。著者らは、そのような環境で好みの楽曲を検索し、再生端末で出しうる最高の品質で再生する、端末適応型音楽エージェントシステムの構築を目標に研究を進めており、本論文ではその原案を述べる。システムは3つの論理的な部分からなる。1.ユーザの要求を表現する形容詞を楽曲の特徴量に変換する部分、2.その特徴量から楽曲を検索する部分、3.再生端末において最適な品質に楽曲を補正・加工する部分、である。本論文ではそれぞれの部分で解決すべき問題について考察する。また著者が行った楽曲の特徴抽出を試みる実験に対する評価も述べる。

キーワード：音楽情報処理、特徴、好み、端末適応、高品質、インターネット

A proposal for a terminal adapted music agent system reflecting user's tastes

Takezawa Naoto[†] Katumoto Michiaki^{††}

[†]The University of Electro Communications ^{††}Communication Research Laboratory

Abstract

Recently more and more music contents are distributed via the Internet. We expect that almost all the music contents will be distributed via the Internet in the near future, and that they will be better quality than CD's. The goal is to build a system that retrieves user's favorite music contents and adapts them to each multimedia player, which can play movies and music contents in order to play them as better quality as possible. In this paper we introduce our system. The system consists of three logical parts: 1.a part converts user's demands described as adjectives to the characteristic parameters of the music contents, 2.a part retrieves music contents with the characteristic parameters calculated by the part 1, 3.a part adapts music contents to each multimedia player so that they play contents as better quality as possible. In this paper we consider some problems to be solved and report on the experiment in taking characteristic parameters out of music contents.

Keywords: music signal processing, characteristics, tastes, terminal adapted, high quality, Internet

1 はじめに

近年楽曲は様々な流通形態で大量に消費者に届けられており、インターネットが発展するに伴い、多くの楽曲が CD 以上の品質でインターネットを通じて流通することが予想できる。個人はその中から自分の好きな楽曲のみを手に入れたいはずであり、またできるだけ良い品質で再生したいはずである。そこで著者らは楽曲の特徴を抽出し、個人の好みを反映した楽曲の検索を行い、それらを再生端末で最適な品質で再生するシステムを構築することを目標に研究を進めている。

楽曲の特徴を抽出し、検索に用いる楽曲検索システムがいくつか報告されている。主旋律の音高差や和音、打楽器の出現頻度などの特徴を利用したもの[1]、頻度のような単純な量ではなく、音楽学の分野で楽曲の構造を解析する際に用いられる「関数的表現」という音のまとまりを単位として楽曲の特徴抽出を行うもの[2]、楽曲の印象を表す形容詞と、リズムなどの楽曲構成要素との対応関係に関する音楽心理学の研究結果を利用し、特徴を抽出するもの[3]などである。

それらはいずれも MIDI 形式の楽曲を解析している。MIDI はコンピュータが演奏する手順を

記述するデータフォーマットであり、MIDI データは楽譜のようなものである。そのそのため MIDI データは音色を完全に表現することが困難である。また、同じ楽譜の楽曲でも演奏する楽器、指揮者の違いで楽曲の印象が異なることがある[4]。ゆえに MIDI データから楽曲の特徴を抽出するのは適切ではないと考える。そこで著者らは録音、あるいは実演されている波形データを解析し、特徴抽出を行う。著者はその1手法として、ウェーブレット変換によって楽曲の特徴を抽出する実験を行った[5]。

さらに、参考文献のいずれの研究においてもユーザの好みは考慮されていない。著者らは、楽曲検索システムはユーザの好みの楽曲を検索できなければ意味がないという立場から、ユーザの「嗜好パラメタ」という概念を導入する。ユーザが気に入っている楽曲の特徴量をそのユーザの「嗜好パラメタ」とする。嗜好パラメタは検索の絞込みに用いられる。

また、参考文献のいずれの研究においても楽曲が再生されたときユーザによる評価を考慮に入れていない。自宅のスピーカシステムで楽曲を鑑賞する場合、高品質な再生を行うためには、高性能なスピーカなどの性能のよい高価な装置を買う必要がある。これらの装置をすべての人が買うことは難しい。そこで、性能が中程度の再生端末でも高性能な端末で再生した音にできるだけ近い品質で再生できるように原音を補正・加工することも目標の1つとして研究を進めている。

2 リファレンススピーカの性能測定実験

「高品質な音」を定義するため、高性能なスピーカ A の周波数特性を測定する。本システムにおける「高品質な音」とは、リファレンススピーカ A と同等の性能を持つスピーカから出力された音とする。

リファレンススピーカ A の音響特性をオーディオ測定装置で測定する。また実験装置のブロック図を図1に示す。

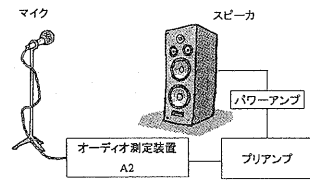


図 1 測定装置のブロック図

2.1 周波数特性の測定方法

リファレンススピーカ A にレベルが 1dBu のピンクノイズを入力し、バンドパスフィルタの通過周波数をスイープさせて周波数特性を測定した。ノイズ発生装置、バンドパスフィルタは測定装置に内蔵されているものを使用した。

2.2 測定結果

リファレンススピーカの周波数特性を図2に示す。実験時の気温は 23°C、湿度は 77%であった。

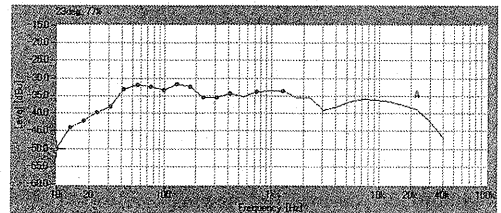


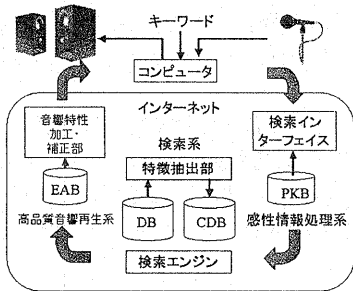
図 2 リファレンススピーカの周波数特性

3 システム概要

システムは大きく分けて 3 つの論理的な部分からなる。(1)感性情報処理系、(2)検索系、(3)高品質音響再生系である。(1)はユーザが要求する雰囲気システムに理解させる部分である。具体的にはユーザが入力したキーワードを対応する楽曲の特徴量に変換する。(2)は(1)で求めた特徴量に適合する特徴を持つ楽曲を検索する部分である。この部分は平行して楽曲の特徴抽出も行う。(3)は必要に応じて再生端末において最適な品質になるように楽曲を補正・加工する部分である。図3にシステムの全体図を示す。また図4にシステムの論理的な構造を示す。

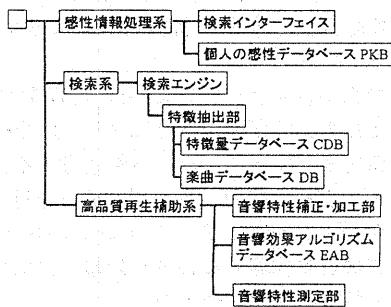
個人の感性データベース (PKB) には個人の「嗜好パラメタ」が保存される。つまり個人の好

みの楽曲に関する情報を入れる予定である。特徴量データベース（CDB）には楽曲の特徴量とアーティスト名、発売元などの情報を保存する予定である。楽曲データベース（DB）にはタイトルなどの情報と楽曲を保存する予定である。音響効果アルゴリズムデータベース（EAB）には音響効果を付加するアルゴリズムを保存する予定である。



PKB: Personal Knowledge DataBase
 CDB: Characteristics DataBase
 EAB: Efficient Algorithm DataBase

図 3 システムの全体図



PKB: Personal Knowledge DataBase
 CDB: Characteristics DataBase
 EAB: Efficient Algorithm DataBase

図 4 システムの論理的な構造

全体的な処理の流れは次のようになる。まずユーザが、聞きたい曲の雰囲気表現していると思われるキーワードを入力する。その要求はインターネット上にある感性情報処理系に送られ、対応するいくつかの特徴量の候補に変換される。それらの内、ユーザの嗜好パラメタに近い特徴量が検

索系に渡される。検索系はあらかじめ楽曲から抽出しておいた特徴量データベースの中から適合する楽曲を検索し、それを高品質音響再生系に渡す。ここでは再生環境の音響特性を参考に、必要に応じて端末において最適な品質になるように楽曲を補正・加工する。特別な要求があれば、音響効果を付加するアルゴリズムを利用し、楽曲に音響効果を付けることができる。再生環境において最適な品質に補正・加工された楽曲がインターネットから配信され、ユーザに届けられる。

4 論理的な部分の詳細

4.1 感性情報処理系

感性情報処理系のブロック図を図 5 に示す。

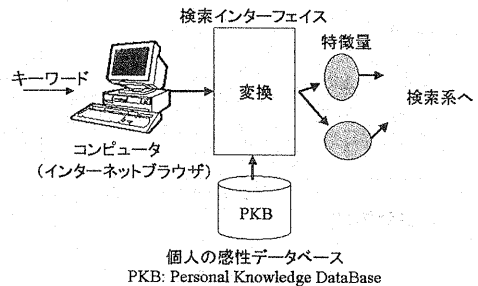


図 5 感性情報処理系のブロック図

感性情報処理系は図 5 に示すように 2 つの部分から構成される。検索インターフェイスと個人の感性データベース（PKB）である。それぞれの機能は次の通りである。

- 検索インターフェイス：キーワードによるユーザの要求をいくつかの楽曲の特徴量に変換する。さらに嗜好パラメタを参考に候補の絞込みを行う。
- PKB：ユーザが指定した好みの楽曲の特徴量を嗜好パラメタとして保存、管理する。

4.2 検索系

検索系のブロック図を図 6 に示す。検索系は図 6 に示すように 4 つの部分からなる。検索エンジン、特徴量抽出部、特徴量データベース（CDB）、

楽曲データベース (DB) である。それらの機能は以下の通りである。

- 検索エンジン: 感性情報処理系で求められた特徴量をもとに CDB から適合する特徴量を検索する。検索の結果得られる楽曲のタイトルなどの情報をもとに、DB から楽曲を検索する。検索した楽曲を高品質音響再生系へ送信する。
- 特徴量抽出部: 信号処理によって楽曲の特徴量とそのタイトルなどの情報を抽出し、CDB に保存、管理する。
- CDB: 特徴量抽出部で得られた楽曲の特徴量とそのタイトルなどの情報を保存、管理する。
- DB: 楽曲そのものとそのタイトルなどの情報を保存、管理する。

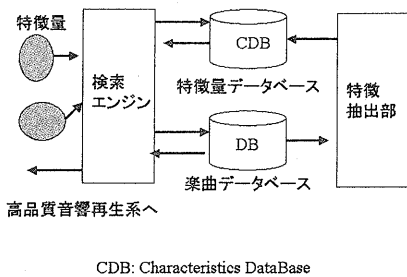


図 6 検索系のブロック図

4.3 高品質音響再生系

高品質音響再生系のブロック図を図 7 に示す。

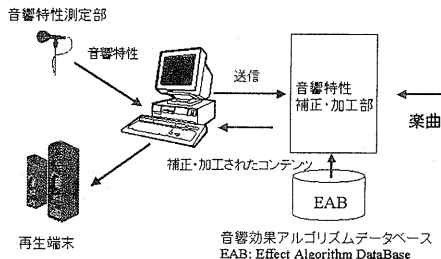


図 7 高品質音響再生系のブロック図

高品質音響再生系は図 7 に示すように 3 つの部分からなる。音響特性補正・加工部、音響特性測定部、音響効果アルゴリズムデータベース (EAB) である。それぞれの機能は以下のようになる。

- 音響特性補正・加工部: 音響特性測定部のデータにもとづき、再生環境において最適な品質になるように検索系で検索された楽曲を必要に応じて補正・加工する。また、特別な要求がある場合、音響効果を付加する。補正・加工されたデータをユーザの端末に送信する。
- 音響特性測定部: 再生環境の周波数特性などを測定し、音響特性補正・加工部へ送信する。
- EAB: ユーザが要求した音響効果を付加するためのアルゴリズムを保存、管理する。

5 解決すべき問題と対策

各論理的な部分で解決すべき問題とその対策について考察する。

5.1 検索インターフェイスの問題と対策

問題は、

- キーワードによる評価をどう評価するか
- その評価値と楽曲の特徴をどのように対応づけるか

の 2 つである。対策としては、音響心理学や感性情報工学の先行研究を参考にする。まず楽曲の雰囲気を表現するキーワードを用意する。キーワードによって楽曲のカテゴリ分けを行う。次にカテゴリに分かれたコンテンツのスペクトル分布などの物理的な特徴を観察し、カテゴリごとの特徴の傾向をつかむ。

5.2 検索エンジンの問題と対策

問題は、要求された特徴量をどこまで厳しく絞りこんで検索するか、である。システムに対する入力も特徴量も曖昧性があるため、厳密な検索を行うことは困難である。これに対しては嗜好パラメタによる絞込みを行う。いくつかの候補の中から、嗜好パラメタに近い特徴量を持つ楽曲を検索する。

5.3 特徴量抽出部の問題と対策

問題は、楽曲のどのような物理的特徴を抽出するか、である。これは検索系で最も重要な問題であり、CDBと共通の問題である。また、この問題は感性情報処理系の問題と密接に結びついている。対策としては、心理学的なアプローチと信号処理的なアプローチの両方から考える必要がある。本論分では信号処理的なアプローチの1手法として、ウェーブレット変換による特徴抽出が有効であると考え、実験を行った。詳細は6章で述べる。

5.4 音響特性補正・加工部の問題と対策

問題は、再生端末において最適な音質に楽曲を補正する必要が生じた場合、再生環境の音響特性に基づきどのような補正を行うか、またどんな音響特性を測定すればよいか、である。これに対しては再生環境の音響特性と音質の心理的な評価の関係を実験により推定する必要がある。そこで特性の異なる数種類のスピーカを用意し、同じ音源に対してどのような音質の違いが現れるかを実験により求める。どのような音響特性が音質に影響があるのかを結果から推定する。

5.5 EABの問題

問題点は、要求された音響効果をどのようにアルゴリズムに変換するか、アルゴリズムをどのように用意するか、などである。

6 楽曲の特徴抽出実験とその評価

著者は、ウェーブレット変換で楽曲の波形データからその特徴を抽出することを試みた。この章ではその実験内容と評価について述べる。

6.1 実音解析による特徴抽出

実音解析の重要性は先に述べた通りであり、音楽は実際に演奏されてはじめて音楽として成立すると考える。著者は楽曲の音色やハーモニーなどの「響き」が、楽曲の雰囲気には大きな影響を与えることに着目した。ここでいう「響き」とは、音を認識するのに十分な時間におけるスペクトルパターンのことをさす。そこで楽曲の波形データをウェーブレット変換した時間-周波数平面のパターンを楽曲の特徴として抽出する方法を検討した。

6.2 ウェーブレット変換

6.2.1 2次元画像への展開手法

波形データを時間-周波数の平面に展開する方法にはフーリエ変換とウェーブレット変換がある。本実験では音の響きの微妙な違いをとらえることが目的であるため、時間分解能の高いウェーブレット変換を用いることにした。音色は波形の時間変化によって特徴付けられる。例えば、周波数が同じで振幅包絡が時間と共に増大する正弦波と、減少するものの2種類の音があるとする。両者は周波数成分が同じだが、前者はオルガンのような、後者はピアノのような音に聞こえることが知られている。このように、ウェーブレット変換はフーリエ変換では得られない情報を得ることが出来る。

6.2.2 基本ウェーブレットの選択

基本ウェーブレットには Gabor を用いた。時間と周波数の不確定性の積が最も小さく、また楽曲のような正弦波を基調とした波形と親和性が高いと判断したためである。Gabor の式を以下に示す。

$$\varphi(t) = \frac{1}{2\sqrt{\pi}\sigma} \exp\left(-\frac{t^2}{\sigma^2} - it\right) \quad (1)$$

ただし i は虚数単位、 σ は時間領域の分散を表す。

6.2.3 解析条件

ウェーブレット解析プログラム GabrIWT を用いて、サンプリングレートが 44.1kHz の wave 形式の部分楽曲データに対して Gabor によるウェーブレット変換を行った。Gabor の分散 σ は 8、解析レベル j は 9、1 オクターブあたりの分割数は 12 とした。

6.2.4 解析結果

楽曲の雰囲気を著者の主観によって「クラシック」と「ポップス」のカテゴリに分け、「クラシック」はさらに「軽やか」と「リラックス」に、「ポップス」は「楽しい」と「リズムカル」に分類した。図 8、図 9 に解析結果の一例を示す。各図の上部には変換を行う前の楽曲の波形データを示してある。色が白いほどエネルギーが高く、黒い部分はほとんどエネルギーが無いことを意味する。 $j=0$ の中心周波数は約 7kHz である。「クラシック」と「ポップス」の画像の間には明確な

違いが現れたが、さらに細かいカテゴリの場合は明確な違いが現れなかった。むしろ変換画像は「クラシック」の中ではバイオリンの曲とピアノの曲、「ポップス」の中では似たような楽器で演奏された曲、というように波形の違いを忠実に表していることがわかった。

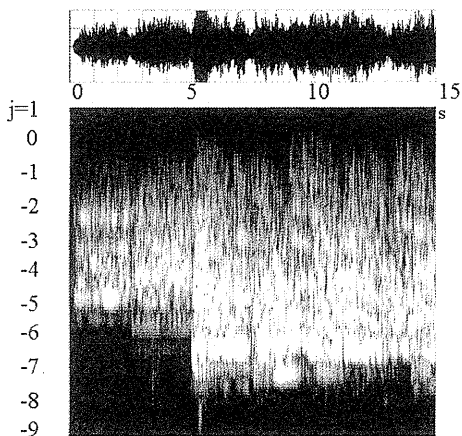


図 8 「クラシック」に属する部分楽曲をウェーブレット変換した時の時間-周波数平面のパターン

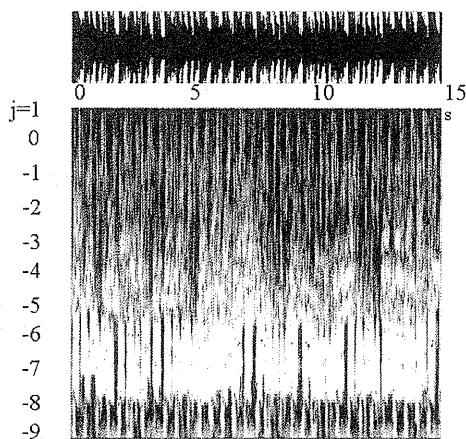


図 9 「ポップス」に属する部分楽曲をウェーブレット変換した時の時間-周波数平面のパターン

6.2.5 実験の評価

実験では部分楽曲をウェーブレット変換により時間-周波数の 2 次元画像に展開することで楽曲の雰囲気やコンピュータ上で表現することを試みた。その結果、十分な特徴抽出は行えなかったが、音色が類似したものどうしの分類には成功

した。これは実音の解析による特徴抽出の可能性を示していると考えられる。楽曲検索システムへ応用するためには、カテゴリを汎用性のあるものを作り直し、さらに実際にスピーカから再生した音を解析して再評価する必要がある。

7 まとめ

本論文では、個人の好みの楽曲を検索し、それを再生端末でより良い品質で再生するように適応させるシステムを提案した。またシステムを構成する部分の機能について述べた。さらにシステム実現にあたり解決すべき課題を挙げ、その対策について考察した。最後に楽曲の特徴抽出の 1 手法を検討し、実験を行った。今後はそれぞれの部分で解決すべき問題、特にユーザが要求する雰囲気と楽曲の特徴の対応関係をどのようにするかという問題を、実験によって解決していく予定である。

参考文献

- [1] 野田 達也, 森山 剛, 小沢 慎治: 印象語による楽曲検索システム, 電子情報通信学会総合大会論文集 (1999)
- [2] 澤戸 周作, 大越 豊, 藤本 克弘, 加藤 宗子, 飯作 俊一: 楽曲知的検索における GA を用いた特徴量抽出法の提案, 1999 年電子情報通信学会情報・システムソサイエティ大会 (1999)
- [3] 吉野太智, 高木秀幸, 清木康, 北川高嗣: 楽曲データを対象としたメタデータ自動生成方式とその意味的連想検索への適用, 情報処理学会研究報告, データベースシステム No.116-41, pp.109-116, (1998)
- [4] Seichiro Namba, Sonoko Kuwano, Tadasu Hatoh, Mariko Kato: Assessment of Musical Performance by Using the Method of Continuous Judgement by Selected Description, Music Perception, Vol.8, No.3, pp.251-276 (1991)
- [5] 竹澤 直人, 高橋 聖, 中村 英夫: ウェーブレット変換を用いた楽曲の特徴抽出に対する一検討, 2000 年電子情報通信学会総合大会論文集 (2000)