

感性に基づく特徴推測による音楽データベース検索システム

三浦 真奈美 三石 大 佐々木 淳 船生 豊

岩手県立大学ソフトウェア情報学部

g031w151@edu.soft.iwate-pu.ac.jp, {takashi, jsasaki, funyu}@soft.iwate-pu.ac.jp

利用者の感性に基づく特徴推測による音楽データベース検索システムを提案する。これは、音楽データベースに対する利用者のアクセス履歴から各利用者の興味の方角性を推測し、利用者の興味に基づき、音楽データの特徴をより詳細に推測し、感性語句によるインデックスを半自動的に生成することで、感性検索を可能にするものである。本手法により、内容を示すインデックスを明示的に作成することが難しい音楽データに対し、効果的な検索を実現することが可能になる。本稿では、本システムの利用実験の結果を示し、本手法およびシステムの有効性について検証する。

A Music Retrieval System which Estimates Characteristics of Data based on Users' Sensibility

Manami MIURA Takashi MITSUISHI Jun SASAKI Yutaka FUNYU

Faculty of Software & Information Science,
Iwate Prefectural University

g031w151@edu.soft.iwate-pu.ac.jp, {takashi, jsasaki, funyu}@soft.iwate-pu.ac.jp

We propose a music retrieval system which could estimate individual characteristics of music data based on users' sensibility. It estimates each user's interest (user model) from the history; which data he/she had accessed, estimates a characteristic of music data (music title model) from history; by which users it had been accessed, and create indices in emotional words from estimated characteristics semi-automatically. However it is difficult to create indices expressing individual characteristics of music data, this system could create more appropriate indices and we could retrieve target data efficiently from music database. In order to show the effectiveness of our system, we perform a simple experiment.

1 はじめに

近年、コンピュータ技術やネットワーク技術の発達に伴い、映像、音楽などのマルチメディアデータが電子化されるようになった。それにより、ネットワーク上にはマルチメディアデータベースが多数存在し、インターネットを介して情報を提供するサービスが展開されている。

そのようなサービスの一例として、音楽データを提供するオンラインミュージックショップが挙げられる。

このようなサービスを効果的に利用するためには、大量のデータを蓄積するデータベースの中から、目的のデータを効率良く検索出来ることが必要になると考えられる。

一般に、音楽データベースの中から音楽データを検索する場合、音楽データに関する付随的な情報を用いる手法が多い。付随的な情報とは、音楽のタイトル、作詞・作曲者名、アーティスト名などが挙げられる。これらの情報を用いた検索は、多くのサービスで利用されており、利用者にも判り易く使いやすい。しかしながら、この検索の場合、利用者が目的の音楽データに関する付随的な情報を把握していなければ、目的のデータを探し出すことはできないという問題点がある。

このような問題を解決する方法の一つとして、利用者の興味や関心、嗜好に基づいた感性検索手法が挙げられる。

そこで我々は、利用者の感性に基づいた検索を可能にする感性検索システムを提案し、実装してきた。本稿では、感性語句による音楽データベース検索システムの概要とシステムの有効性を検証するために行った評価実験について述べる。

2 関連研究

本章では、本提案手法およびシステムと関連する既存の研究を挙げ、その問題点について述べる。

利用者の興味や関心に応じた感性検索を行うためには、データそのものの内容に対する特徴付けを行い、その特徴と感性語句を対応付けてインデックスを作成し、それを基に検索を行える必要がある。

絵画や画像データベースにおいて、検索対象データの色や形状、デザインパターンなどを解析して特徴量を抽出し、その特徴量と感性語句を対応付けて、インデックスを作成することで感性検索を可能にする方法がある [1][2]。

しかしながら、音楽においては、曲調からは明るい印象を受けるにも関わらず、歌詞からは暗く悲しい印象を受けるといった複雑な印象を与えるものが存在する。このような音楽に対して、メロディや音高、テンポなどの楽譜データを解析するだけでは、その内容の方向性を掴む

ことは難しい。

音楽データベースを対象にした研究において、辻らの研究では、音高と音長の局所パターンを特徴量としており、予め判っている特定のパターンのみ特徴付けが可能である [3]。また、局所パターンを用いているため、特徴量が音楽の内容の方向性を表しているとは言い難い。

佐藤らの研究では、試聴実験の評価と楽譜データの解析結果の関係から、特徴量を求める変換式を導き出し、その式を用いて特徴量を求め、対応付けを行っている [4]。吉野らの研究では、予め求められている印象語の相関量を引用し、楽譜データの解析結果から、数学モデルを用いて特徴量を求めている [5]。これらの方法では、変換式や数学モデルを用いることにより、予め判っている特定のパターン以外の未知のパターンを持つ音楽にも特徴付けが可能である。しかしながら、いずれも、クラシック音楽などを対象としており、歌詞を含むような、複雑な印象を与える音楽に対して有効であるとは限らない。

また、利用者の興味や関心に基づいたデータの分類や情報提供を行うために、利用者のデータベースに対するアクセス履歴を利用する手法がある [6][7]。これらの手法では、アクセス履歴から、利用者がどのような語句で検索を行い、検索結果から得たデータはどのような語句で特徴付けが行われているかといった分析を行っている。これにより、利用者の好みを推測して興味や関心を反映した検索結果を提示することができ、また、データのそのものの内容の方向性を推測することが出来る。

しかしながら、これらの手法は、検索や情報提供時に不要なデータを除去するためのフィルタリングや、利用者の嗜好に合ったデータのみを提供するリコメンドが目的であり、そのままでは、感性検索のためのインデックスを作成することは困難である。

3 感性に基づいた特徴推測

本章では、我々の提案してきたアクセス履歴を利用した特徴推測手法について説明する。

3.1 アクセス履歴を利用した特徴推測

我々はこれまで、音楽データのように複雑で曖昧な特徴を併せ持ち、明示的なインデックスを作成困難なマルチメディアデータベースに対し、利用者のデータベースアクセス履歴に基づくインデックスの半自動生成手法を提案し、[8][9]. 音楽データベースへの適用を検討してきた[10].

本手法は、音楽のようなデータの場合、利用者の興味や関心に応じてデータの利用に偏りが生じる事が予想され、データベースのアクセス履歴を観察することで、その利用傾向から、データが潜在的に持つ方向性を推測出来ると考えるものである。

例えば、「明るい」という感性語句により表現される特徴を持つ音楽データをよく利用する利用者が存在した場合、その利用者は「明るい」音楽を好むのではないかと推測される。また、このような「明るい」音楽を好む利用者から多く利用される音楽データは、「明るい」特徴を少なからず持っているとして推測することが出来る。

このように、利用者のアクセス履歴を観察することにより利用者の興味の方向性(ユーザモデル)を推測し、このユーザモデルから、音楽データが潜在的に持つ内容の方向性(音楽タイトルモデル)を推測することが可能であると考えられる。この推測結果をもとに、個々の音楽データについて各感性語句ごとに重み付けを行うことで、その特徴を示すインデックスを半自動的に生成することが可能となる。

3.2 各モデルのベクトル表現

感性検索におけるインデックスを作成するために、3.1節で示した各音楽データの音楽タイトルモデルおよび利用者のユーザモデルを、各軸に感性語句を当てはめた n 次元のベクトル空間により表現する。

音楽データの集合 $\{m\}$ および利用者の集合 $\{n\}$ が存在すると仮定し、音楽データ m の推測される特徴、および利用者 n の推測される興味の方向性をそれぞれ各成分が0から1の大きさを持つベクトル空間により表現し、これを音楽タイトルモデル $\vec{T}_m(c_m)$ 、ユーザモデル $\vec{U}_n(c_n)$ と定義する。ただし c_m および c_n は、それぞれ音楽データ m が利用者から利用された回数、および利用者 n がなんらかのデータを利用した回数を示す。例えば仮に、ここでは音楽タイトルモデルおよびユーザモデルのそれぞれが感性語句 $K1$ から $K6$ までの6次元のベクトル空間により表現出来るものとする。このとき、 $K1$ を特徴として多く持ち、 $K2$ から $K6$ は特徴としてあまり持たない音楽データ m の音楽タイトルモデルの初期値、すなわち $c_m = 0$ における音楽タイトルモデル $\vec{T}_m(0)$ は、

$$\vec{T}_m(0) = (1.0, 0.0, 0.0, 0.0, 0.0, 0.0)$$

のように表現することが出来る。

また同様に、利用回数 c_n において推測されているユーザモデル $\vec{U}_n(c_n)$ を

$$\vec{U}_n(c_n) = (0.8, 0.2, 0.3, 0.5, 0.1, 0.4)$$

のように表現することが出来る。

このように、音楽データの内容の特徴、および利用者の興味の方向性を共通の次元を持つベクトルで表現することにより、予め大まかに推測された音楽タイトルモデルから、アクセス履歴に基づきユーザモデルを推測し、同様にユーザモデルから、より詳細な音楽タイトルモデルを推測することが可能となる。また、この音楽タイトルモデルから、各感性語句に対応する重み付きインデックスを生成することが出来る。

ただし、ここでは全ての感性語句が直交しているものとする。

3.3 特徴推測の手順と各モデルの更新式

3.3.1 特徴推測の手順

本システムにおいて、本手法による特徴推測は以下の手順で行う。(1) 個々の音楽データに対

して大まかな特徴推測を行い、音楽タイトルモデルを作成する。また、ユーザモデルは初期値を設定する。(2) 各利用者が選択した音楽データの音楽タイトルモデルの集合を基に、利用者のユーザモデルを推測し、更新する。(3) 各音楽データを利用した利用者のユーザモデルの集合を基に、音楽データの音楽タイトルモデルを推測し、更新する。

データベースの利用時に(2)、(3)を繰り返すことにより、利用者の興味の方角を推測し、また、この利用者の興味に基づき、音楽データの複雑で曖昧な特徴を表現することが出来る。

3.3.2 各モデルの更新式

音楽タイトルモデルおよびユーザモデルの更新には、推測された音楽タイトルモデルとユーザモデル間の相互フィードバックに基づく更新式を用いる。更新式は、例えば、(1)、(2)のように定義することが出来る。ただし、音楽データ m 、利用者 n 、時刻 t であり、利用者 n が t までに利用したデータの個数を c_{U_n} および音楽データ m を t までに利用した利用者数を c_{T_m} と表す。

$$\bar{T}_m(c_{T_m} + 1) = \frac{c_{T_m} \times \bar{T}_m(c_{T_m}) + \bar{U}_n(c_{U_n})}{c_{T_m} + 1} \quad (1)$$

$$\bar{U}_n(c_{U_n} + 1) = \frac{c_{U_n} \times \bar{U}_n(c_{U_n}) + \bar{T}_m(c_{T_m})}{c_{U_n} + 1} \quad (2)$$

これは、音楽データ1回の利用毎に、利用者 n がそれまでに利用した音楽データの音楽タイトルモデルの平均、および音楽データ m をそれまで利用した利用者のユーザモデルの平均に基づき、それぞれユーザモデルおよび音楽タイトルモデルを更新することを意味する。

4 システムの設計と実装

本章では、3章で述べた特徴推測手法を用いた、感性語句による音楽データベース検索システムの設計、およびそのシステムの実装について説明する。

4.1 システムの概要

本システムでは、利用者毎にユーザモデルを作成するため、先ず最初にログインをして専用の環境を入手する。検索画面から感性語句を選択し入力すると、その感性語句を特徴として持つ音楽データのタイトルリストがその特徴の大きさに応じて特徴付けされて表示される。そのタイトルリストから、音楽データを選択すると、データの詳細情報を参照でき、実際に音楽データを試聴することが出来る。

本システムは、WWW上のオンラインミュージックショップへの適用を想定しているため、利用者が音楽データを選択した時点では決定とせず、一時的にリストに追加し、最終的に購入を決めた時点で音楽データが決定したとして、各モデルの更新を行うこととした。

4.2 システム構成

感性語句による音楽データベース検索システムを図1図2に示す形で設計した。

本システムは、音楽検索エンジン、音楽データベース、およびユーザインタフェースを提供するクライアントから構成される。

音楽検索エンジンは、検索機能部、特徴推測部の2つからなり、音楽タイトルモデルベースおよびユーザモデルベースを有する。

音楽タイトルモデルベースは、各音楽データのID、タイトル、アーティスト名、および音楽タイトルモデルを持ち、同様に、ユーザモデルベースは、各利用者のIDとユーザモデルを持つ。

また音楽データベースには、音楽データの詳細情報および、データそのものを格納する。

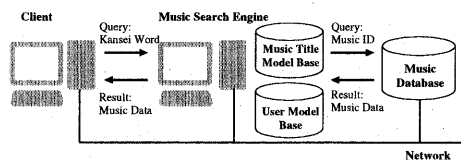


図1: システム構成

音楽検索エンジンにおいて、検索機能部は、クライアントが提供するユーザインタフェースを介して利用者とのインタラクションを行い、

クライアントからの検索要求に応じた音楽タイトルモデル、および音楽データベースの検索を行い、その結果の提示をする。また特徴推測部は、利用者のアクセス履歴から、各モデルの更新式を用いて、音楽タイトルモデルおよびユーザモデルの更新を行う。

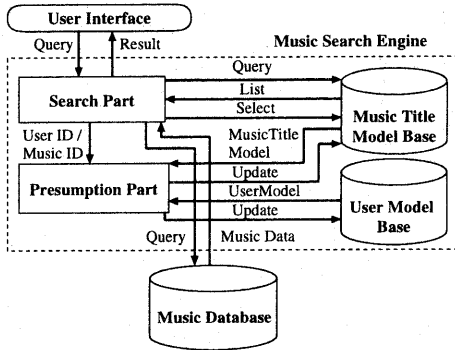


図 2: 音楽検索エンジンのアーキテクチャ

4.3 システムの実装

4.2 節で述べた設計に基づき、感性語句による音楽データベース検索システムの実装を行った。実装環境を表 1 に示す。

また、音楽データベースには、無作為に選出した日本のアーティストの作品 500 曲の音楽データおよび情報を格納した。音楽データは、全て MP3 形式で保存した。

表 1: システムの実装環境

OS:	Solaris 2.7
DBMS:	PostgreSQL 7.0.2
Web サーバ:	Apache 1.3.12
記述言語:	Java 1.2
ライブラリ:	servlet, sql 等
servlet エンジン:	Tomcat 3.1

この 500 曲に対して、筆者の 1 名が試聴しながら各感性語句に対して特徴がどの程度あるか分析し、0 から 1 の間で 0.1 刻みで値を付け、この値について、0.5 以上を 1、それ以下を 0 とおおまかに振り分け、音楽タイトルモデルの初期値を作成した。

ユーザモデルの初期値は各感性語句とも 0.5 とした。

感性検索および各モデルの表現に用いる感性語句は、文献 [3][11] を参考に、表 2 に示す 24 語を選択した。

実装したシステムの実行例を図 3 に示す。

表 2: 感性語句一覧

明るい	楽しい	元気な	陽気な
暗い	悲しい	気が沈む	重々しい
厳かな	壮大な	神秘的な	哀れな
寂しい	静かな	優しい	穏やかな
優雅な	素朴な	野生的な	神秘的な
激しい	情熱的な	華やかな	騒がしい

5 評価実験

4 章で述べた感性語句による音楽データベース検索システムの利用実験を行った。提案手法による、音楽タイトルモデルおよびユーザモデルの特徴推測を行い、その有効性を検証する。

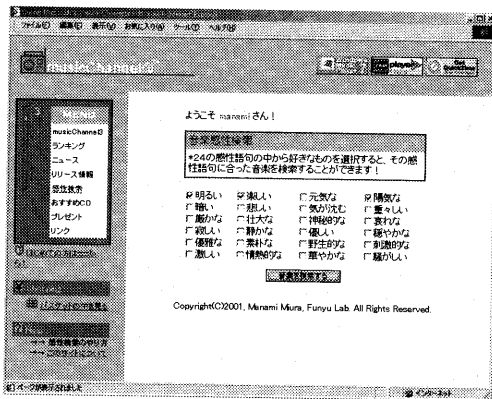
5.1 実験内容

本システムの利用実験を行い、音楽タイトルモデルおよびユーザモデルの特徴推測を行った。

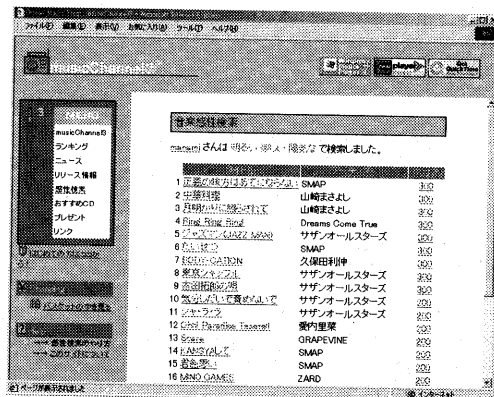
実験に参加する被験者は、無作為に選出した、岩手県立大学ソフトウェア情報学部の学生 10 名である。また、音楽データ、音楽タイトルモデルおよびユーザモデルの値は、4.3 節で述べた内容と同じものを用いた。実験の手順は以下の通りである。

被験者の興味、嗜好に基づき、感性語句を選択して検索を行い、検索結果から任意の音楽データ 1 曲を選択してリストに追加してもらう。この動作を繰り返し 10 曲選択したところで、決定ボタンを押し音楽データを購入する。これを 1 セットとし、合計 10 セット行った。

システム側で、購入処理と同時に、音楽タイトルモデルとユーザモデルの更新を行い、更新前と後に分けてログを記録する。ログの内容は、



(a) 検索画面



(b) 検索結果の提示

図 3: 感性検索システムの実例

各モデルの更新時刻、アクセス対象の音楽データの ID、利用者の ID、検索に用いた感性語句、音楽タイトルモデルまたはユーザモデルである。このログを解析して、音楽タイトルモデルの値を抽出する。

5.2 評価方法

提案手法およびシステムの有効性を示すために、実験結果の評価を行う。

音楽タイトルモデルの初期値は、音楽データを人手によって詳しく分析して付けられた値を元に、大まかに特徴付けされている。よって、提案する特徴推測手法により、大まかな値から、より詳細な値に更新されていけば、本手法は有効であると考えられる。

以上の考えから、ここでは、複数の被験者によって分析された音楽データの分析結果(以下これを分析値と呼ぶ)を正解として、特徴推測が行われる度に、音楽タイトルモデルが分析値の平均に近づいているか(以下これを相違度と呼ぶ)を調べる。

また、特徴推測における各モデルの表現および検索に利用した感性語句に関しては検証を行っていないため、分析値を作成する際、個人の感性語句に対する意味の捉え方の違いを考慮する必要があると考えられる。そこで、より正

確に評価を行うために、分析値の付け方について、各被験者に大きな違いがあるかを感性語句ごとに検証し、違いのない感性語句を求め、これらの語句の分析値を用いて、相違度を求めることとする。

これにより、評価は以下の手順で行う。1. 被験者間で分析値に大きな差のない感性語句を求める。2. 1の結果を用いて相違度を求める。

手順 1. 被験者間で分析値に大きな差のない感性語句を求める。

(1) 複数の被験者に音楽データを分析してもらい、分析値を作成する。

被験者は、実験に参加した 9 名である。被験者に分析してもらう音楽データは、コストの問題から、全 500 曲中、比較的利用回数が多く、且つ複数の利用者から利用されているデータ 10 曲とした(以下サンプルデータと呼ぶ)。分析値の作成は、4.3 節で述べた分析方法と同様に行った。

(2) 被験者間で分析値の付け方に、大きな差のない感性語句を求める。

各被験者の分析値の差の比較には、対応のある K 群の処理間の差の検定であるフリードマン検定を用い、感性語句ごとに合計 24 回行う。検定は、帰無仮説 H_0 : 「被験者間に差はない」、対立仮説 H_1 : 「被験者間に差はある」として、有意水準 $\alpha = 5\%$ の条件で行い、有意確率 P を

求める。 $P > \alpha$ の時、帰無仮説が採択され、被験者間に差のない感性語句が求められる。

手順 2. 1 の結果を用いて相違度を求める。

(1) 手順 1 において求められた被験者間に差のない感性語句ごとに、分析値の平均を求める。

分析値の平均は、正確性を重視するため、最大値と最小値を除外して求めることとする。

(2) 各感性語句の分析値の平均と音楽タイトルモデルから相違度を求める。

相違度は、(3) 式により定義される。ただし、ここでは、相違度 d 、分析値の平均 \bar{A} 、音楽データ m 、音楽タイトルモデル T 、感性語句 k 、感性語句の総数 n と定義し、音楽データ m の音楽タイトルモデル T は、 T_m と表す。また、分析値の平均 \bar{A} の、 i 番目の感性語句 k を $\bar{A}(k_i)$ と表し、同様に、音楽タイトルモデル T_m の i 番目の感性語句 k を $T_m(k_i)$ と表す。

$$d = \frac{\sum_{i=0}^n |\bar{A}(k_i) - T_m(k_i)|}{n} \quad (3)$$

相違度は、分析値の平均と音楽タイトルモデルの差を意味するものであり、特徴推測の回数が増加するに従い、値が減少すれば分析値の平均に近づいていると考える。

5.3 結果

被験者の分析結果から分析値を得た。

この分析値に基づいて、被験者間に差のない感性語句を求めた。フリードマン検定の結果、各被験者の分析値の付け方に差のない感性語句は、明るい、元気な、陽気な、寂しい、野性的な、激しい、刺激的な、情熱的な、華やかな、騒がしいの 10 語であると判った。

この 10 語の感性語句について、分析値の平均を算出し、この分析値の平均と音楽タイトルモデルから、(3) 式を用いて相違度を求めた。この相違度の経過を図 4 に示す。

この結果、サンプルデータ 10 曲中、特徴推測の回数が増加するに従い、相違度が減少したデータは 8 曲、逆に相違度が増加したデータは 2 曲であった。10 曲中 8 曲のデータの相違度が減少し、また、平均 0.154 程度減少しているこ

とから、概ね良い結果を得ることができたといえる。

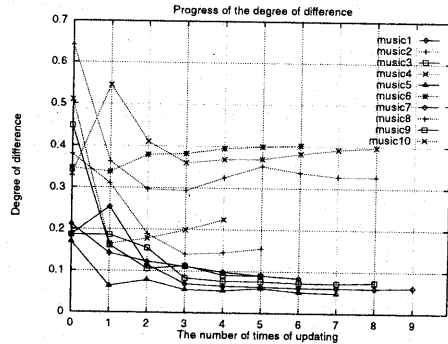


図 4: 相違度の経過

しかし、逆に相違度が増加してしまったデータもあり、この理由としては、初期値の傾向が、分析値の傾向と大きく異なっていたことが考えられる。また、更新式が音楽タイトルモデルとユーザモデル間の相互フィードバックに基づくものであるため、音楽タイトルモデルの特徴推測がうまくいかなかったことが考えられる。例えば、音楽データの分析値の傾向と大きく異なる興味の方性を持つ利用者が、このデータに対して一時的に興味を持ちアクセスしたことで、音楽タイトルモデルの値が大きく変わってしまった可能性が挙げられる。

以上の結果から、本手法による特徴推測は概ね有効であり、よって、本システムは有効であるといえるが、音楽タイトルモデルの初期値の設定方法、各モデルの更新式について、改善が必要であると考えられる。

6 おわりに

本稿では、感性に基づく特徴推測による音楽データベース検索システムについて提案した。

本システムは、利用者の興味や関心に応じてデータの利用に偏りがあることに着目し、データベースのアクセス履歴から利用者のユーザモデルを推測し、これに基づき、個々の音楽データの音楽タイトルモデルを推測することによって、複雑で曖昧な特徴を持つ音楽データのインデックスを半自動的に生成し、感性検索を可能にするものである。

また、本稿では、本システムの有効性を示すために、実際にシステムを利用した評価実験を行った。この実験の結果、本手法による特徴推測は概ね有効であり、よって、本システムも有効であることが判った。

しかしながら、初期値の特徴推測への影響や初期値を作成する際のコストの問題、音楽タイトルモデルの不適切な更新などの問題があり、これらを改善していく必要がある。

また、今回の評価実験では、音楽タイトルモデルの更新回数が最大で9回と少なかったため、今後より長期的な評価実験を行い、本手法およびシステムの検証を行う予定である。

参考文献

- [1] 原田将治, 伊藤幸宏, 中谷広正: 感性語句を含む自然言語文による画像検索のための形状特徴空間の構築, 情報処理学会論文誌 Vol. 39, No. 4, pp. 2356-2366 (1999).
- [2] Fukuda, M., Sugita, K. and Shibata, Y.: Perceptual Retrieving Method for Distributed Design Image Database System, *Trans. IPS Japan*, Vol. 39, No. 2, pp. 158-169 (1998).
- [3] 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研究報告, SP, 音声, Vol. 96 Num. 565 pp. 17-24 (1997).
- [4] 佐藤聡, 菊地幸平, 北上始: 音楽データを対象としたイメージ検索のための感情価の自動生成, 情報処理学会研究報告 DBS, Vol. 99, No. 39, pp. 57-64 (1999).
- [5] 吉野太智, 高木秀幸, 清水康, 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, DBS, Vol. 98 No. 116 pp. 109-116 (1998).
- [6] 風間一洋, 佐藤進也, 清水奨, 神林隆: WWWのユーザ操作履歴によるHTML文書の相関関係の解析, 情報処理学会論文誌, Vol. 40, No. 5, pp. 2450-2459 (1999).
- [7] 橘高博行, 佐藤直之, 鈴木英明, 曾根岡昭直: パーソナライズ情報提供方式の提案と評価, 情報処理学会論文誌, Vol. 40, No. 1, pp. 175-187 (1999).
- [8] 三石大, 佐々木淳, 船生豊: ユーザの利用履歴を利用した動的なインデックス半自動生成手法の提案, 情報処理学会研究報告 2000-DBS-121, Vol.2000, No.44, pp.53-60(2000).
- [9] Mitsuishi, T., Sasaki, J. and Funyu, Y.: A Proposal of Semi-automatic Indexing Algorithm for Multi-media Database with Users' Sensibility, *Proc. of the 2000 Spring Conference of KOSES & International Sensibility Ergonomics Symposium*, pp. 120-125 (2000).
- [10] 三浦真奈美, 三石大, 佐々木淳, 船生豊: 感性語句による音楽データベース検索システムの構築, 第62回情報処理学会全国大会講演論文集(3), pp. 69-70 (2001).
- [11] 谷口高士: 音楽と感情, 北大路書房, (1998).