m-RIK：個人の感性特性に対応可能な音楽検索システム

杉 原 太 郎[†,☆]　森 本 一 成[†]　黒 川 隆 夫[†]

個人の感性特性に対応可能な音楽検索システムを提案するとともに，個人特性の学習結果や検索結果に関する評価実験について述べる．本システムは学習によって個人の感性を標準的な感性に変換するニューラルネットワーク（NN）と，検索キーとして与えられる感性値にマッチする楽曲を提示する音楽データベース（MD）より成る．個人の特定楽曲についての感性を 40 の感性語対に関する主観評価値を要素とする感性ベクトルで表し，多数の楽曲の感性ベクトルを配置した感性空間により個人の感性を特徴づける．NN は個人の感性空間から標準の感性空間への変換法を学習するもので，希望する楽曲のイメージを感性ベクトルとしてシステムに入力すれば，MD の中からそれにマッチするものが出力される．感性空間と検索について述べた後，学習実験を行い，正しく感性空間の学習が行われることを示した．また実験者が与えた感性ベクトルをキーとして検索した楽曲に対して主観評価実験を行った結果，MD に入れた楽曲が少なかったにもかかわらず被験者の 74%が検索結果に適合すると回答した．さらに，個人の感性特性を学習させずに検索を行うシステムと性能比較実験を行った結果，m-RIK の成績が高いことを確認した．

m-RIK: Music Retrieval System Specialized in Individual Kansei Characteristic

Taro Sugihara,[†,☆] Kazunari Morimoto[†] and Takao Kurokawa[†]

We proposed a music retrieval system on the basis of an individual kansei trait, and its learning of individual kansei traits and an experiment for evaluating its music retrieval performance were described. The system consists of a neural network (NN) that becomes to transform any individual's kansei characteristics to a standard one through learning and a musical database (MD) that presents tunes matching to a key kansei value. A kansei vector whose elements were subjective evaluation ratings on 40 kansei word pairs represented an individual's kansei to a specific tune, and his/her kansei trait was characterized by a kansei space that had his/her kansei vectors corresponding to many tunes. The results of an experiment showed that NNs could currently learn transformation from individual kansei spaces to the standard one. Although many tunes were not contained in the MD, 74% of the subjects were suited with the results of the system's retrieval. It was also shown that the m-RIK was more useful than the system without the NN.

1. はじめに

まだ聴いたことのない楽曲の中から自分が望むものを探しだす場合，大きく 2 つの方法が考えられる．1 つは過去の経験から蓄積した知識に頼ってジャンル，作曲者，演奏者などの名義情報をキーとして絞り込む方法である．考えたキーに一致する名義情報を持つ楽曲で満足できる場合は別として，この方法では知識外の項目名義を持つ楽曲にはアクセスできない．また，候補の楽曲が得られても，それが聴きたいと思うイメージに一致するかどうかは実際に曲を聴くまで確信が持てないなどの欠点がある．もう 1 つの方法は欲する曲のイメージを「力強い」,「さわやかな」など複数の感性語によって検索するものである．この方法では高い確度で自分のイメージに合う楽曲が得られると期待できる．Kim ら[1] は，音楽の専門家ではない 22 名の被験者に 7 曲のクラシックを検索させた際に,「emotions」（全クエリの 24%で使用）に分類された単語が,「occasions or events」に分類された単語（全クエリの 29%で使用）と並んで多用されたことを報告しており，このことからも感性語による検索の重要性が示唆されている．第 1 の方法で音楽検索を行うのは容易で，すでに多くのシステムが実用化されている．しかし後者の感性音楽検索を行えるシステムはまだ研

† 京都工芸繊維大学大学院工芸科学研究科
Graduate School of Science and Technology, Kyoto Institute of Technology
☆ 現在，北陸先端科学技術大学院大学
Presently with Japan Advanced Institute of Science and Technology

究途上のものである．

感性語をキーとする感性検索法には，絵画や画像を検索対象にしたもの[2)〜6)]や服飾製品を検索するもの[7)]などがある．また，検索結果を求めることが主目的ではないものの，個人の感性空間や概念空間を利用して情報を提示する研究には，ユーザの興味や知識に応じて個人化された情報提供を行うシステム Takealook[8)]や，感性語の入力に応じて個人特性を反映した住居のデザイン候補を提示するシステム Lipstick[9)]などがある．

一方，感性音楽検索システムに関する研究には北川らのシステム[10)]や池添らのもの[11)]がある．北川らは楽曲のメディアデータの自動抽出を行い，そのデータに人の感性を反映した関数を導入してメタデータを作成し，それを用いる検索法を提案した．池添らのシステムは「音楽感性空間生成部」と「検索システム部」から成る．「音楽感性空間生成部」では，楽曲の印象に対して行った因子分析の結果を基に構成した音楽感性空間を検索空間とし，遺伝的アルゴリズムと重回帰分析を使用して未知の楽曲の自動インデクシングを可能とした．そして音楽感性空間に配置された楽曲の座標値と検索したい楽曲の座標の予測値に基づいてユーザの希望する楽曲を提供する．これらのシステムは音楽の物理特性とユーザの感性特性を対応づけて検索を行うものであるが，扱える音楽の物理量は単純なもの（Standard MIDI File: SMF）に限定されている．

本研究では感性の個人特性に対応可能な音楽検索システム（A Music Retrieval System Specialized in Individual Kansei Characteristic: m-RIK）を提案する．感性には個人差があり，音楽の場合も例外ではない[12)]．本論文では個人の感性特性を表現するために，40 種類の感性語対を導入し，各感性語対について楽曲の 7 段階主観評価を行わせ，その結果を感性空間と呼ぶ 40 次元空間の点で表す．複数の楽曲がこの空間でどのように分布するかによりユーザの感性特性を表すことができる．本システムはニューラルネットワーク（NN）と音楽データベース（MD）より成る．まず NN が各ユーザの感性空間から標準となる特定個人（Standard Evaluator: SE）の感性空間へのマッピングを学習する．次にこの結果に基づいて，MD 内の楽曲を評価して配置してある SE の感性空間からユーザが 40 個の感性値で表した欲する楽曲を探索し，提示する．

以下，2 章で感性語対の選択と感性空間について，3 章で提案するシステムについて述べる．さらに個人の感性特性の学習と感性検索に関するシミュレーションおよびその結果を 4 章で紹介し，5 章においてシステムの評価実験について述べ，6 章で提案システムの問題点などの議論を行う．

2. 感性空間による個人の感性特性の表現

2.1 感性語の収集および選択

個々の楽曲に対してどのようなイメージをいだくかは個人によってかなり異なり，その結果として曲ごとの好き嫌いが生じる．したがって楽曲に関するイメージを定量的に記述できれば，個人の音楽に関する感性特性を表現でき，それを基に個人に特化した音楽検索システムが構築可能になると期待される．音楽に関する研究では楽曲のイメージを表すのに適した感性語対を求め，各感性語対について個々の楽曲にいだくイメージを SD（意味微分）法[13)]で定量化することにより，特定個人が各楽曲に対して持つイメージを総合的に表現することにする．

評価に用いる感性語を決定するために，多数の感性語を収集し，そのうちから楽曲のイメージを表現するのに適した感性語対を求めることにした．まず多数の感性語を集めたものとして定評のある日本カラーデザイン研究所[15)]のデータを採用した．このデータはカラーイメージを表現する感性語を集めたものであるが，基本的な感性語のほとんどが含まれているうえ，音楽や他のイメージ表現にも利用できる語句を多く含んでいる．さらに音楽のイメージ表現に適した感性語と最近の若年者が用いる感性語を補う目的で，比較的若い読者層を持つ音楽情報雑誌「ROCKIN' ON JAPAN」約 2 年分のバックナンバ（1998 年 9 月号〜2000 年 12 月号）を対象に，CD・音楽ビデオのレビュー欄から名詞を修飾している 339 語を抽出した．両データから同一語句を削除すると 530 語が得られ，これを本研究で用いる感性語対を求めるための原データとした．

次に音楽の印象をイメージしやすい語句を調査するために，この感性語群を用いてアンケート調査を行った．具体的には「あなたは次の表に書かれている単語から「音楽」をイメージできますか？」との問いに対して，各語ごとに「イメージ可」,「イメージ不可」，また提示した感性語の意味が分からなければ「意味不明」のいずれかをチェックさせた．アンケートの有効回答者は 10 代後半〜20 代後半の男女 149 名（男性：81 名，女性：68 名）であった．

アンケート調査の結果から,「意味不明」の回答率が 1/3 を超えるもの,「イメージ不可」の回答率が 1/2 を超えるものを除き，残った語句について国語辞典，類義語辞典および英和辞典を用いて正確な意味を調べ,

表 1　楽曲イメージ評価用の感性語対
Table 1 Kansei word pairs for evaluating tune image.

番号	感性語対	番号	感性語対	番号	感性語対
1	静かな — にぎやかな	15	緊張した — リラックスした	29	冬らしい — 夏らしい
2	落ち着いた — 勢いのある	16	平凡な — 刺激的な	30	俗っぽい — 神秘的な
3	泣ける — 笑える	17	不透明な — 透き通った	31	せかせかした — ゆったりした
4	ありきたりな — 新鮮な	18	眠たくなるような — 目の覚めるような	32	不快な — 心地よい
5	さめた — 熱い	19	素朴な — 飾り気のある	33	なつかしい — 目新しい
6	日常的な — ドラマティックな	20	伝統的な — 革新的な	34	リアルな — ロマンティックな
7	知的な — ワイルドな	21	しんみりした — うきうきした	35	弱々しい — 力強い
8	スローテンポの — アップテンポの	22	おとなしい — 元気な	36	悲しい — 楽しい
9	うっとうしい — さわやかな	23	冷たい — 温かな	37	重厚な — 軽快な
10	軽量感のある — 重量感のある	24	ちまちました — 壮大な	38	陰気な — 陽気な
11	秋らしい — 春らしい	25	暗い — 明るい	39	生気のない — 生き生きした
12	人工的な — 自然な	26	単調な — メリハリのある	40	いらいらした — 安らぐ
13	ドライな — センチメンタルな	27	無感動な — 感動的な		
14	止まっているような — ダイナミックな	28	穏やかな — 激しい		

類義語群のうちから「イメージ可」の回答率の高かった語句を選択して，251 の感性語に絞った．

音楽の印象を SD 法により計測するには，意味が両極になる感性語対が必要であるので，このため反意語が存在しない語句を除外した☆．そして得られた感性語の中から「イメージ可」の回答率が高く，反意語を含まない上位 40 語を採用することにした．最後に反対語辞典により各語句の反意語を 1 つ定め，楽曲のイメージを評価するための感性語対を**表 1** のように 40 種類作成した．

2.2 感 性 空 間

これまでに提案された音楽感性空間には，SD 法を用いて音楽聴取実験を行い因子分析の結果得られた因子を軸としたもの[11),14)] がある．これらの手法は，被験者全体の印象を空間構築のためデータとして使用しており，個人の感性特性は考慮されていない．本研究では SD 法により得た楽曲の評価データから，各人の感性空間を定義する．

2.1 節で得られた 40 の感性語対を用いてユーザ i が楽曲 j を 7 段階 SD 法で評定した結果を v_{ijk}（$k=1,\ldots,40$）とする．たとえば，表 1 の「静かな—にぎやかな」（$k=1$）に 7 段階の評価尺度を与え，静かな：1，どちらでもない：4，にぎやかな：7 という評価値を割り当てる．1〜4 および 4〜7 の尺度間には程度を表現せず，程度の解釈を被験者の判断に任せた．v_{ijk} を要素とする式 (1) の 40 次元ベクトル V_{ij} をユーザ i の楽曲 j に対する感性ベクトルと呼ぶ．

$$V_{ij} = (v_{ij1}, v_{ij2}, \ldots, v_{ij40}) \tag{1}$$

人間の感性は多様であり，同じ楽曲を聴いたときにいだくイメージに個人差のあることが知られている[12)]．したがって n 個の楽曲に対する感性ベクトルの感性空間における分布によってユーザ i の感性特性を表すことができる．以下では楽曲 j と感性ベクトル V_{ij}（$j=1,\ldots,n$）を対応させた 40 次元空間をユーザ i の感性空間と呼び，$\boldsymbol{V_i}$ と表記することにする．定義により感性空間は離散的であるが，便宜上これをユークリッド空間と見なす．

ユーザとは別に，音楽鑑賞力を有し，広いジャンルの楽曲に対して一貫した評価を行うことのできる人物 SE の感性空間を $\boldsymbol{V_{\mathrm{SE}}}$ と表す．5 章で述べるように，$\boldsymbol{V_{\mathrm{SE}}}$ には上述の n をはるかに超える数 n_{SE} の楽曲に対する感性ベクトルを配置させる．本研究では楽曲の検索を任意の感性ベクトルをキーとして行うことを考えるが，このために NN を用いて $\boldsymbol{V_i}$ の V_{ij} から $\boldsymbol{V_{\mathrm{SE}}}$ の V_{SEj}（$j=1,\ldots,n$）への写像を学習させる．これを単に感性空間 $\boldsymbol{V_i}$ の学習と呼ぶことにする．

3. m-RIK

提案するシステムの概要を**図 1** に示す．システムは 1 名のユーザの感性特性を学習して，SE の感性特性に変換する NN と，SE による評価によってあらかじめ感性ベクトルが付与されている多数の楽曲を含む音楽データベースより成る．楽曲を検索する場合，ユーザは検索したい楽曲のイメージを感性ベクトルで表し，それをシステムに入力する．システムの NN はこの感性ベクトルを SE の感性ベクトルに変換し，MD はこのベクトルと登録された各楽曲の感性ベクトルのユークリッド距離を計算し，距離の近い楽曲を検索結果として出力する．

このような検索が可能となるためには，まずユーザ i の感性を NN に学習させる必要がある．必要なデータとして，ユーザに少数（n とする）の楽曲を聴かせ，表 1 の感性語対について 2.2 節で述べたように 7 段階評価をさせて得られた n 個の感性ベクトル V_{ij}（$j=1,\ldots,n$）を用いる．教師データは SE に同様の評価をさせた結果 V_{SEj}（$j=1,\ldots,n$）である．

図 2 にシステムで使用した NN の構成図を示す．NN

☆「イメージ可」の回答率が高く，反意語が存在しない語句には「切ない」や「ファンキーな」などがあった．

は，感性ベクトルの要素に対応する 40 個のニューロンを持つ入力層と 40 個のサブネットワークで構成されている．1 つのサブネットワークは各感性語対に対応しており，それぞれが学習機能を持つ 40 個のニューロンから構成された 1 層の中間層，SE の感性に変換した結果を感性ベクトルとして出力する 1 個のニューロンを持つ出力層からできている．NN がユーザ i の感性特性を正しく学習した場合，入力感性ベクトル V_{ij} に対する NN の出力 $O_{jk} = (o_{j1}, o_{j2}, \ldots, o_{j40})$ は SE の感性ベクトル V_{SEj} と一致する．k 番目の出力ニューロンは V_{SEj} の第 k 要素 v_{SEjk} に対応する値 O_{jk} を出力する．学習アルゴリズムにはバックプロパゲーション法を採用する．

一方，MD は学習が終了して以降，入力された任意の感性ベクトルが SE の感性ベクトルに変換された結果を用いて，MD に納められている多数の楽曲の感性ベクトルとのユークリッド距離を計算し，値が小さい楽曲を複数個ユーザに提示する．以後，システムが学習を行う時期を学習相，検索に利用する時期を検索相と呼ぶ．

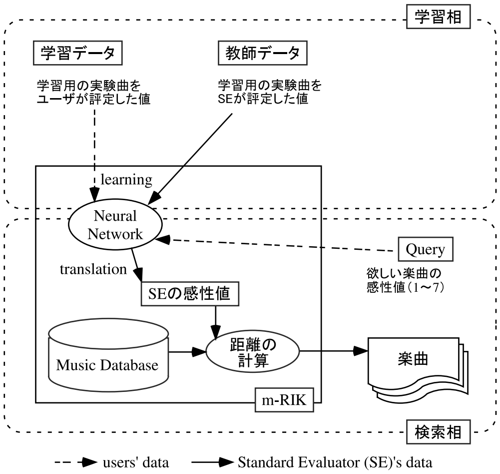


図 1　m-RIK の概要
Fig. 1　Overview of m-RIK.

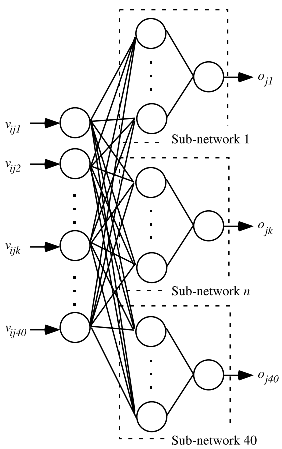


図 2　NN の構成
Fig. 2　Construction of NN.

学習が完了した後の検索相で確度の高い検索が行われるためには，多数の楽曲が一貫した感性によって評価されていることが必要で，本研究ではこの役割を SE に委ねたことになる．SE がよく似たイメージを持つと判断した曲は，感性空間 $\boldsymbol{V}_{\mathbf{SE}}$ で近くに位置することになる．

提案するシステムを使用して検索する際，40 個すべての感性語対とその強度を選択する方法と，任意のもののみ指定する方法が考えられる．任意のもののみ指定する方法のほうがユーザの負担は少なくなることは自明であるが，入力される感性語対が減少することがシステムにどのような影響を及ぼすのかが明らかになっていないため，本研究では 40 個すべての感性語対とその強度を選択する方法を採用した．

4. 個人の感性特性の学習

4.1　個人の感性特性の計測

個人の感性特性をシステムに学習させるデータを収集するために，音楽聴取実験を行って個人ごとの感性特性を計測した．本論文は前章で述べたシステムの有効性を示すことを主な目的としているので，実験に用いる楽曲は小学校の授業に使用される鑑賞曲に限定した．東京書籍発行の「新訂　新しい音楽」（平成 12〜13 年度用）の小学 1 年生から 6 年生までの指導書で取り上げられている多数の鑑賞曲のうちより，小学校の音楽教諭と相談して，曲調（調や拍子など）ができるだけ多様になるように**表 2** に示す 12 曲を選定した．

防音が施された実験室で，CD に録音したこの 12 曲を BOSE 社の AMS-DMC によって各被験者に聴か

表 2　感性計測用の 12 曲
Table 2　Twelve tunes used for measuring individual kansei characteristics.

No.	曲名
1	ブラジル
2	子ぞうの行進
3	メヌエット長調
4	管弦楽のための木挽き歌 「山で木挽きの仕事に合わせて歌われる木挽き歌」
5	シルクロード
6	風の丘「魔女の宅急便」より
7	ハンガリー舞曲第5番
8	トルコ行進曲(オーケストラ)
9	ジャマイカン　ルンバ
10	コンドルは飛んで行く
11	きらきら星変奏曲　KV.265
12	オラリー

表 3 学習結果の 1 例

Table 3 An example of the learning results.

感性ベクトル V_{i1}			感性ベクトル V_{i2}			感性ベクトル V_{i3}			感性ベクトル V_{i4}			感性ベクトル V_{i5}			感性ベクトル V_{i6}		
No.	tunes	distance	No.	tunes	distance	NO.	tunes	distance	No.	tunes	distance	No.	tunes	distance	No.	tunes	distance
1	ブラジル	0.78	2	子ぞうの行進	0.68	3	メヌエット	0.97	4	木挽き歌	1.08	5	シルクロード	1.18	6	風の丘	0.90
2	子ぞうの行進	2.20	1	ブラジル	2.25	11	きらきら星	3.70	6	風の丘	6.27	6	風の丘	3.34	5	シルクロード	3.71
9	ジャマイカ	3.07	8	トルコ行進曲	3.80	12	オラリー	4.02	10	コンドル	6.49	12	オラリー	3.56	10	コンドル	3.86

感性ベクトル V_{i7}			感性ベクトル V_{i8}			感性ベクトル V_{i9}			感性ベクトル V_{i10}			感性ベクトル V_{i11}			感性ベクトル V_{i12}		
No.	tunes	distance	No.	tunes	distance	NO.	tunes	distance	No.	tunes	distance	No.	tunes	distance	No.	tunes	distance
7	ハンガリー	1.09	8	トルコ行進曲	0.98	9	ジャマイカ	1.01	10	コンドル	0.94	11	きらきら星	1.02	12	オラリー	1.13
8	トルコ行進曲	3.35	1	ブラジル	3.35	1	ブラジル	2.77	5	シルクロード	3.62	12	オラリー	2.25	11	きらきら星	2.47
1	ブラジル	4.28	7	ハンガリー	3.55	2	子ぞうの行進	3.27	6	風の丘	3.67	3	メヌエット	3.70	5	シルクロード	3.33

せた．被験者は 20 代の学生の男性 14 名であった．この被験者のうち，以前の実験[16] で最も成績の良かったもの 1 名を SE とした．

実験では楽曲を初めて聴く被験者と，以前に聴いたことがある被験者との経験の差が印象に与える影響を少なくするために，同じ曲を 2 回聴かせた．1 回目は曲全体の印象を把握するために 1 曲通して聴かせ，2 回目に聴きながら曲の印象を 7 段階で評定させた．曲の提示順序は被験者ごとにランダムとした．この実験より，13 名分の 12 曲に関する感性ベクトル V_{ij}（$i=1,\ldots,13,\ j=1,\ldots,12$）および V_{SEj}（$j=1,\ldots,12$）が得られたことになる．

4.2 感性特性の学習とその結果

前述の SE の感性ベクトル V_{SEj} を教師データとし，残りの 13 名の V_{ij} を学習データに用いて 40 個の NN に学習を行わせた．なお NN の内部では感性ベクトルの要素値 7〜1 を 1〜0 に規格化し，学習の収束条件を NN の各出力と教師データの平均 2 乗誤差が規格化した値で 0.005 以下になることとした．また慣性項の重み更新係数 α は 0.8，速度項の重み更新係数 η は 0.2 とした．収束までの学習回数には被験者間で差があったが，すべての被験者について学習過程は収束した．

次に学習が正しく行えたかどうかの検証を行った．**表 3** は 1 名の被験者についての学習結果であり，12 個の入力感性ベクトルごとに MD でマッチングのとれた楽曲を距離の短い順に 3 曲ずつあげた．距離の計算は，NN の 0〜1 の出力値を 1〜7 の評定値に戻して行った．この被験者が「ブラジル」に付与した感性ベクトル V_{i1} を学習後の NN に入力したところ，NN は「ブラジル」を出力した．同様に，残りの 11 の感性ベクトルについても正しい楽曲を出力できた．この表よりすべての感性ベクトルについて最も距離の短いものとして正しい楽曲が出力されていること，また 1 位の曲と 2 位以下の曲との距離がかなり大きいことが分かる．すなわち，学習した楽曲については，NN がこの被験者の感性空間をかなり正確に SE の感性空間に変換できるようになった．他の被験者についても同様の解析を行ったところ，全被験者かつ 12 曲すべてにおいて正しい楽曲が最も近い曲として出力されたことを確認した．さらに，1 位の曲と 2 位以下の曲との距離が大きい点も同様であった．

表 4 汎化能力の評価結果

Table 4 Result of the leave-one-out cross validation.

曲名	ブラジル	子ぞうの行進	メヌエット	木挽き歌
相関	0.70	0.54	0.60	0.14

曲名	シルクロード	風の丘	ハンガリー	トルコ行進曲
相関	0.73	0.55	0.39	0.48

曲名	ジャマイカ	コンドル	きらきら星	オラリー
相関	0.52	0.65	0.63	0.76

NN の汎化能力を調べるため，leave-one-out cross validation を行った．被験者 1 名を選び，表 2 の中から「ブラジル」を除いた 11 曲を用いて感性特性を学習させ，上記の方法と同様に「ブラジル」の感性ベクトル V_{i1} を学習後の NN に入力し，各サブネットワークの出力値と SE の「ブラジル」の感性ベクトル間の Peason の相関係数を求めた．残り 11 曲，全被験者に同様の行程を繰り返し，各曲の相関係数の平均値を求めた．これを示したものが **表 4** である．「木挽き歌」を除く 11 曲については，相関係数が 0.4 以上とまずまずの値であった．「木挽き歌」は 0.14 と非常に低い値であるが，これは「木挽き歌」が他の曲と大きく異なる音楽的性質を有しており，感性空間で近くに配置された楽曲が存在しなかったためである．「木挽き歌」に対する被験者の評定を見ても，「いらいらした」（平均評定値：2.4，残りの 11 曲の平均評定値：4.9），「陰気な」（平均評定値：1.9，残りの 11 曲の平均評定値：4.8）など，他の曲には与えられてない評定値が付与されていた．このことから，学習時に使用した楽曲とかけ離れていない印象を与える楽曲に対しては汎化能力を持つことが判明した．

5. システム評価実験

5.1 実験の手続き

本システムの有効性を確認するために，主観評価実験を行った．

まず 4.1 節で選定した SE に新たに 57 の器楽曲を聴かせ，その印象を評定させた．これらの曲は前述の

表 5 評価実験に使用した感性ベクトル
Table 5 Kansei vectors for the system evaluation in the retrieval phase.

感性ベクトル＼感性語対	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
A	1	1	4	4	4	6	1	1	6	4	4	4	6	4	5	4	4	4	4	2
B	7	7	4	4	7	4	6	7	4	4	4	4	4	4	4	4	4	4	4	4
C	4	3	3	2	4	4	4	2	3	3	3	4	4	2	3	4	4	3	4	4
D	6	4	6	4	4	4	4	4	7	1	7	4	4	5	5	4	4	5	4	4

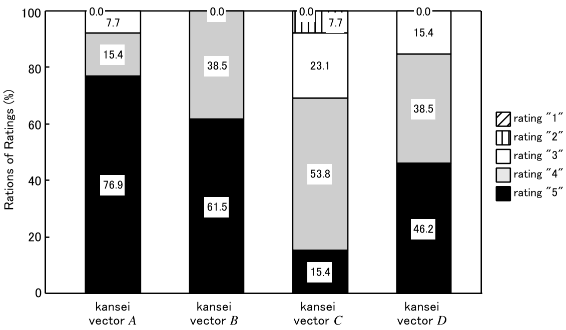
感性ベクトル＼感性語対	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
A	1	1	4	4	4	4	7	1	4	4	7	7	4	7	4	4	4	4	4	7
B	7	7	5	4	7	5	4	6	4	4	3	5	4	4	7	7	4	7	6	5
C	3	3	3	3	2	4	3	3	3	4	4	3	4	4	4	2	3	3	4	3
D	7	5	6	4	6	6	5	5	5	4	4	6	4	4	5	7	7	7	7	6

小学生用教科書の指導書の鑑賞曲 38 曲を基調に，ロックやその他の器楽曲を 19 曲加えたものである．次にこの 57 曲および実験曲 12 曲，計 69 曲の SE による評価結果を感性ベクトルとして MD に格納し，SE を除く 13 名に m-RIK の検索結果を主観評価させた．検索に使用した感性ベクトルは全被験者に共通で，**表 5** に示す 4 種 ***A***，***B***，***C***，***D*** である．この 4 ベクトルは感性空間で離れるように，特に ***C*** は他と大きく異なる位置にくるように決定した．感性ベクトル ***A*** は「静かな」や「落ち着いた」に，***B*** は ***A*** と反対に「にぎやかな」や「勢いのある」，***D*** は「さわやかな」,「春らしい」,「軽快な」に重みをおいて構成した．***C*** には「スローテンポな」や「暗い」などの感性語に重みを置き，他のベクトルと大きく異なるプロフィールを与えた．m-RIK は 4.2 節で各被験者の感性特性を学習した 13 名分の m-RIK を用いた．まず 4 個の感性ベクトルを各 m-RIK に入力し，MD 内の楽曲中からそれぞれの感性ベクトルとのユークリッド距離が短い上位楽曲 5 曲を求めて，その m-RIK に対応する被験者に提示した．そして紙面で用意した入力感性ベクトルを見せながら，検索された楽曲がどの程度マッチングしているかを 5 段階（1：まったく合っていない，2：あまり合っていない，3：どちらともいえない，4：ある程度合っている，5：非常に合っている）で，ベクトルごとに主観評価させた．最後に，システムに対するコメントを自由記述式のアンケートにより求めた．

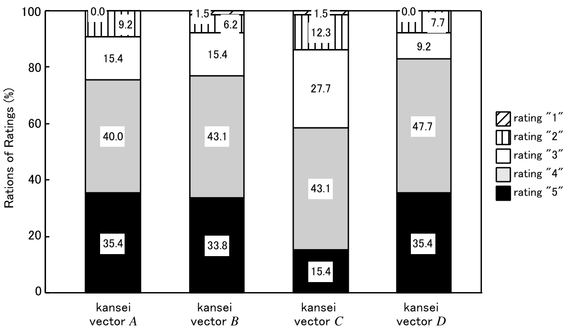
5.2 実 験 結 果

主観評価実験の結果を**図 3** に示す．検索キーとして用いた 4 種類の感性ベクトルごとに，(a) には検索された 1 位の曲に対して全被験者が判断した評価値の割合を，また (b) には 5 曲すべてに対する同様の結果を百分率で表した．

まず入力感性ベクトルに最も近かった曲 (a) に関して，感性ベクトル ***A***，***B*** および ***D*** の 3 ベクトルの場合，検索結果がクエリと合っていたと回答した（評定値 4 と 5）被験者の割合（適合度）は 84.6%〜100%と非常に高かった．さらに，合っていなかったと回答した被験者（評定値 1 と 2）の割合はいずれも 0%であり，この 3 ベクトルについては良好な結果を得たといえる．感性ベクトル ***C*** は 5.1 節で述べたように小学



(a) 提示順が1位の楽曲に対する主観評価



(b) 検索された5曲に対する主観評価

図 3 主観評価実験結果
Fig. 3 Results of the experiment for subjective evaluation of system performance.

生用の鑑賞曲にはマッチしにくいプロフィールを与えていたためか，3 ベクトルと比較すると劣った．それでも，合っていたと回答した被験者の割合は 69.2%と高い値であった．検索された全 5 曲に対する評価 (b) は (a) と比較すると全体的に下がったものの，適合度は十分高かった．***A***，***B*** および ***D*** の場合，適合度（75.4〜83.1%）は不適合度（7.7〜9.2%）のおよそ 8〜10 倍であった．ただし，この 3 ベクトルでは評定値 5 の割合が大きく減っており，***A***（30.4%減少）と ***B***（22.7%減少）の減少幅が大きかった．「非常に合っている」の評価が減少したことは，今後，出力曲の距離を計算方法の検討や，距離そのものに基準を設ける必要性を示唆する．***C*** では評定値 4 の割合が 10%減少し，評定値 1〜3 の割合がそれぞれ増加した．

MD に納めた楽曲が 69 曲にすぎず，検索キーとし

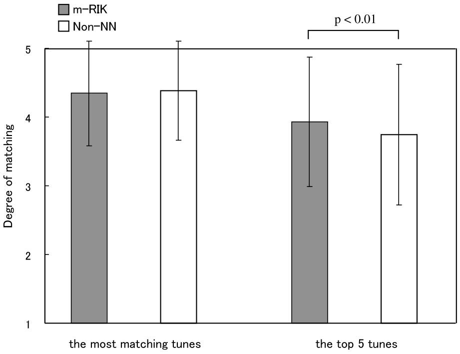


図 4 学習の有無による主観評価の差
Fig. 4 Comparison of system performance between the m-RIK and the system without the NN.

た感性ベクトルから被験者がイメージするものに適合する曲に近いものが含まれている可能性が高くなかったことと，検索用感性ベクトルが実験者から与えられたものであり，被験者がそれから楽曲をイメージすることが必ずしも容易ではなかったことを考え合わせると，被験者の評価は高かったといえる．

さらに，個人特性を学習したことが検索結果に影響を与えるかどうかを調べるために，NN を介さない，つまり学習相を経ないシステムを構築した（以下，Non-NN と呼ぶ）．そして，これまでの実験と同一の被験者を対象に，***A*** から ***D*** の 4 ベクトルを入力し，出力された 5 曲がどの程度マッチングしているかを 5 段階で主観評価させた．その比較結果を**図 4** に示す．

入力感性ベクトルに最も近かった曲に対する評価に対しては，m-RIK（mean = 4.35，S.D. = 0.76）と Non-NN（mean = 4.38，S.D. = 0.72）ともに高い成績を得た．t 検定を行った結果，両者に有意差は認められなかった（両側検定：$t(51) = -0.299$，$p = 0.76$）．続いて，全 5 曲に対して同様の解析を行った．m-RIK が mean = 3.93，S.D. = 0.94 で，Non-NN が mean = 3.74，S.D. = 1.02 であった．t 検定の結果，両者には有意差が認められた（両側検定：$t(259) = 2.65$，$p < 0.01$）．

m-RIK と Non-NN の 1 位の検索曲で同一の楽曲は全 52 曲（13 人 × 1 曲 × 4 ベクトル）中に 1 曲も存在しなかった．Non-NN の出力曲はいずれの被験者にも共通であるため，Non-NN の 1 位の楽曲については本研究で選んだ SE と全被験者の感性評価が近かったものと考えられる．この現象がこの SE 特有のものなのかどうかについては，今後検討しなければならない大きな課題である．全 5 曲について見ると，全 260 曲（13 人 × 5 曲 × 4 ベクトル）のうち，172 曲が異なる提示曲であった．この 172 曲の平均評定値を比較すると，m-RIK が mean = 3.90，S.D. = 0.94 で，Non-NN が mean = 3.67，S.D. = 1.08 であった．このことから，両者で入れ替わった楽曲への評価の差が全 5 曲で有意差が生じた要因であると考えられる．

最も近い楽曲の平均評価値が 4.4 と高いこと，全 5 曲に対しては Non-NN との間に有意差が認められたことから，m-RIK により個人特性を学習したことの有用性が示せたといえる．

検索キーである感性ベクトルと検索された楽曲のマッチングを見るために，被験者 3 名の検索結果を**表 6** に示した．検索キーとした感性ベクトルに最も近い楽曲でも距離は 2.0 前後であった．この距離は個人の感性空間の学習結果と比較すると約 2〜3 倍となっており，1 位の曲であっても検索キーに非常にマッチしたものとはいえない．いずれの感性ベクトル間にも特別な違いは認められない．他の被験者についても同様であった．これは学習用感性空間の精度，および MD 内の楽曲数の少なさによるものと考えられる．学習用楽曲を多くして汎化能力を向上させることや，より多くの楽曲を MD に登録することにより，検索キーとマッチした楽曲が提供できるようになるため，検索結果に対する適合度の向上が期待できる．

検索結果の個人差について検討するため，表 6 下欄に被験者 1 と残りの 2 名の間で異なる楽曲が提示された割合を示した．***A*** ではいずれの被験者も 5 曲中 4 曲が，***D*** においても 5 曲中 4 曲が被験者 1 への提示曲とは異なったものであった．***B*** では各被験者間で楽曲の入れ替わりが少なく，被験者 2 では 1 曲，被験者 3 では 2 曲のみが被験者 1 への提示曲と違っていた．***C*** においては被験者 1 と被験者 2 の間で提示された楽曲群はやや似かよっている（40%が入れ替わった楽曲）が，被験者 3 については異なった楽曲が提示された（60%が入れ替わった楽曲）結果となった．同様の解析を全被験者に対して行い，その平均を計算したところ，***A*** では 53.3%，***B*** では 36.9%，***C*** では 48.7%，***D*** では 53.6%，全体としては 48.1%の提示楽曲が被験者間で異なっていた．まったく同じクエリを入力したにもかかわらずこれほどまでに多様な楽曲が提示されるということは，本システムで提案した感性の個人性を考慮した検索システムを開発することの必要性が示唆されたといえる．

評価実験で用いた SE の感性空間における 69 曲の配置と SE の感性空間へ写像した検索用の感性ベクトルおよび学習用楽曲の関係を視覚化するために，多次元尺度法（Multidimensional Scaling: MDS）によって次元の削減を行った．ユークリッド距離モデルを採

表 6 検索結果の一部

Table 6 Part of the results of music retrieval.

(a) 感性ベクトル***A***に対する結果

Order of matching	Specialized in subject 1		Specialized in subject 2		Specialized in subject 3	
	tunes	distance	tunes	distance	tunes	distance
1	MD−051	2.83	MD−012	2.33	MD−012	2.69
2	MD−059	3.08	MD−004	2.38	MD−015	3.02
3	MD−066	3.13	MD−036	2.38	MD−029	3.03
4	MD−057	3.38	MD−055	2.40	MD−058	3.08
5	MD−012	3.42	MD−034	2.41	MD−034	3.09
turnover of tunes (%)			80		80	

(b) 感性ベクトル***B***に対する結果

Order of matching	Specialized in subject 1		Specialized in subject 2		Specialized in subject 3	
	tunes	distance	tunes	distance	tunes	distance
1	MD−058	2.34	MD−058	1.98	MD−058	2.59
2	MD−059	2.86	MD−023	2.33	MD−023	3.05
3	MD−051	3.06	MD−059	2.49	MD−031	3.13
4	MD−066	3.11	MD−051	2.61	MD−014	3.22
5	MD−023	3.12	MD−015	2.73	MD−059	3.28
turnover of tunes (%)			20		40	

(c) 感性ベクトル***C***に対する結果

Order of matching	Specialized in subject 1		Specialized in subject 2		Specialized in subject 3	
	tunes	distance	tunes	distance	tunes	distance
1	MD−008	2.49	MD−008	2.30	MD−018	2.57
2	MD−027	2.52	MD−019	2.37	MD−008	2.84
3	MD−035	3.13	MD−027	2.52	MD−011	2.90
4	MD−010	3.17	MD−010	2.69	MD−027	2.97
5	MD−040	3.28	MD−047	2.75	MD−062	3.00
turnover of tunes (%)			40		60	

(d) 感性ベクトル***D***に対する結果

Order of matching	Specialized in subject 1		Specialized in subject 2		Specialized in subject 3	
	tunes	distance	tunes	distance	tunes	distance
1	MD−051	2.83	MD−012	2.33	MD−012	2.69
2	MD−059	3.08	MD−004	2.38	MD−015	3.02
3	MD−066	3.13	MD−036	2.38	MD−029	3.03
4	MD−057	3.38	MD−055	2.40	MD−058	3.08
5	MD−012	3.42	MD−034	2.41	MD−034	3.09
turnover of tunes (%)			80		80	

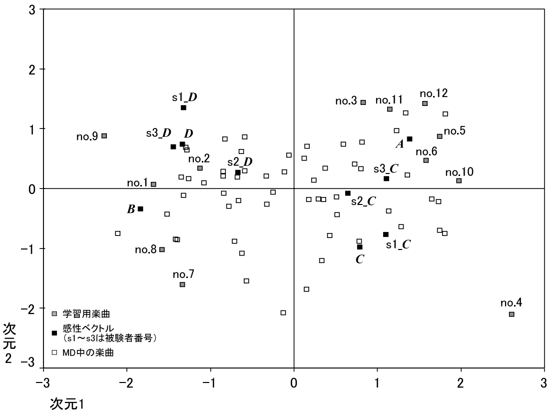


図 5 MDS により描かれた SE の感性空間

Fig. 5 Kansei space of the SE in MDS.

用し，収束値 0.01 で反復推定を行ったところ，最終的にモデルの適合度を表す RSQ は 0.924 となった．次元数を 3 以上にしても Stress 値に大きな改善が見られなかったため，次元数を 2 とした．**図 5** はこの 2 次元に削減した SE の感性空間内に，□は MD に格納した楽曲の配置を，また■は感性ベクトルを表したものである．感性ベクトルについては，表 5 の ***A*** から ***D*** の 4 ベクトル，および感性ベクトル ***C*** と ***D*** について表 6 の被験者 1（s1）〜3（s3）用に SE の感性ベクトルに翻訳した後の配置を示した．この図から，各被験者の翻訳後の感性ベクトルは相互に離れて布置されていることが読み取れ，本システムで感性の個人差を学習・写像できたことが分かる．

続いて，学習用 12 曲に関して検討する．本研究では鑑賞曲の曲調が多様になるように，鑑賞曲群の中から小学校の音楽教諭に選び出させた．図 5 より，学習用楽曲は SE の感性空間上で散らばって布置されていることが読み取れ，様々な特徴を持った楽曲群を学習用に選曲できたことが示された．しかしながら，「木挽き歌」（No.4）や「ハンガリー舞曲第 5 番」（No.7），

表 7 m-RIK に対する被験者の意見

Table 7 Subjects' opinion for m-RIK.

記号	システムに対する意見	意見の数
A1	「どちらでもない」の扱いに困った	4
A2	評価しやすい項目としにくい項目がある	4
A3	合っている項目より合っていない項目のほうが気になる	2
A4	どのくらいの項目が合っていれば「非常に合っている」なのか，判断が難しい	2
A5	自分（被験者）が検索するときには使用することがないものが使われていた	2
B1	曲の好き嫌いが評価に影響する	4
B2	曲調が変化するものの評価が難しい	2
	その他	13

「トルコ行進曲」（No.8）など空間上に偏って配置された楽曲も存在するため，今後学習用楽曲の選曲に関してまだまだ検討する余地がある．4.2 節で示したように，学習していないような楽曲に対しては汎化能力が大きく低下するためである．

表 7 はシステムに対するコメントの中から，システムに対する問題点を項目としてまとめたものである．複数の被験者からあげられた問題点のみをピックアップした．項目は大きく分けて感性語に関するもの（A1〜A4）と，楽曲に関するもの（B1〜B2）に分けられた．アンケートの中で最も多かった回答は A1，A2 および B1 であり，いずれも評価のしやすさに関するものであった．

6. 検　　討

個人の感性特性を構築するための曲を小学生用鑑賞曲から選んだこと，MD に収めた楽曲の数が少なかったこと，検索キーとなる感性ベクトルを実験者が指定したことが制約になっているにもかかわらず，検索結果についての評価は比較的高かった．本章では，今後改善すべき点について議論する．

6.1 感性空間について

感性特性を構築するために各人に評価させる楽曲は，可能な限り感性空間全体に散らばって配置されるよう選ぶことが望ましい．5 章で用いた感性ベクトル ***C***

に対して得られたような結果はこれによって避けることができると思われる．図5を見ると，今回の選曲は右上方に偏っていたが，この一部を他の曲調のものと入れ替えれば，感性ベクトル $\boldsymbol{C}$ のような楽曲にも対応可能な感性空間を構築できる可能性がある．学習用の空間を見てみると，特にNo.4とNo.7の間やNo.4とNo.10の間，No.2とNo.3の間を埋めるような楽曲を組み込む必要がある．曲調だけではなく，曲数についても言及しておかなくてはならない．本論文では12曲を採用したが，感性特性を明確にするには曲数を多くする必要がある．しかし曲数を増やすことはユーザの負担になるので，両者のトレードオフの検討が今後必要となる．

本論文では小学生用鑑賞曲を対象としてシステムを構築したが，実用化を考えればより多くのジャンルの楽曲を検索できるようにしなければならない．対象となる曲のジャンルを拡大する方法はいくつか考えられる．

(1) すべてのジャンルについて1個の感性空間，1個のMDを使用する方法．この方法は検索結果に関するユーザの一致度を低下させると思われる．好みのジャンルは個人により大きく異なり，興味の持てない多数の楽曲を評価させられることはユーザにとって非常に重い負担となるうえ，好みのジャンル以外からも楽曲が検索されてしまうためS/N比は悪くなってしまう．そのため，本当に欲しい楽曲を得るまでに強いられる手間が多くなることが予想される．

(2) 1個の感性空間，ジャンル別にMDを用意する方法．この方法では希望するジャンル（複数選択可能）について感性空間を構築する．最も現実的で，ユーザの指定するジャンルから検索されるので十分な満足が得られると考えられる．日常的に聴く楽曲のジャンルが固定されている人向きである．しかし自分の好みを他のジャンルにも拡張したいと考えている人にはもの足りない．ジャンル別にMDを用意するというのはユーザから見た場合であって，楽曲にジャンルのタグを付けておけばシステムのMDは1個でよい．

最後に，感性空間の次元数について検討する．感性空間を何次元とすれば必要十分かという問題は難しい課題である．今回は40の感性語対を選んだが，多数の楽曲を相互に区別するには次元が高いほどよい．しかし表7のA1，A2のように被験者の多くが評価しにくい項目や強度についての問題を指摘している．感性空間の構築と検索に際してこれだけ多数の項目で評価するのは実際的ではない．また本研究では楽曲のイメージを表現しやすい感性語対を選んだが，評価のしやすさという観点を盛り込む必要があると思われる．項目の種類と数を変えて感性特性の表現に適した感性空間を構築する方法を調べたい．

6.2 SEについて

4章では十分な音楽鑑賞力を持つ被験者をSEとしたが，誰をSEに選んでもシステム構築は可能だろうか．良いSEを選んだm-RIKとそうでないm-RIKでは，おそらくその性能に差が出るであろうと予測される．

ここで著者らが考える良いSEに求められる資質を2点あげておく．

(1) あらゆる楽曲に対してそれらの差異を聞き分けることができ，かつ詳細な評価を行える者．楽曲間の差異を感性空間に反映できなかった場合は，SEの感性空間そのものが妥当性を欠いたものになる．

(2) 中長期的な期間にわたって，感性評価が安定している者．

しかし，どのようにすればこの条件を満たしているSEを選ぶことができるのかは，まだ分からない．こうした条件を満たすSEを見つけ出す方法を考案することも本研究における重要な課題として残されている．

MDには多量の楽曲を収容する必要があるが，これらの曲の感性評価を特定のSEに依存することは非常に難しい．4章での感性特性の学習からはSEの負担を低減する手法が示唆される．本来のSE以外にあらかじめ複数の SE_l（$l = 1, \ldots, m$）を選定しておき，SE_l の感性特性からSEの感性特性への変換法を4章と同じ方法で学習しておく．MDに加えるべき楽曲が生じた場合には，SEあるいは SE_l のうちの1名がこれを感性評価し，SE_l の場合ならばこの評価をSEの感性ベクトルに変換してMDに加えればよい．SE，SE_l の選定は慎重に行う必要があるが，本システムを実用化するために不可欠の手法であると考える．

6.3 検索キーとしての感性ベクトルについて

検索に際して未聴取の楽曲に対する感性ベクトルのすべての要素を確定することは不可能なことも多く，実際的でもない．ある感性語対についてはどの評価値であってもかまわないという場合もありうる．このような場合，ベクトルの要素に “don’t care” を指定できるようにしておき，評価値1〜7のすべてについて検索を行って，検索の和集合から上位楽曲を出力する方法が解決策の1つとして考えられる．

7. おわりに

本論文では個人の感性特性に対応可能な音楽検索システムを提案した．そして個人特性の学習結果およびシステムの評価実験から，

- 全ユーザの感性特性が個人差を反映して学習できたこと
- 提示される楽曲のパターンがユーザごとに異なること
- その楽曲に対する適合度が 74%（全 5 曲）であること

という 3 つの結果が得られ，システムとしての有用性を示すことができた．本システムに用いた手法は SE やユーザの感性特性が比較的長期にわたって安定していることを前提としており，同じ楽曲に対する評価が少し時間を置くと変化するようなユーザには不向きである．

今後，6 章で述べた課題についても検討し，システムの実用化に向けてユーザの一致度の向上を図る予定である．また，提案した手法は音楽以外の感性検索にも適用可能と考えている．

謝辞 本研究において，実験の選曲に貴重なご助言をくださった京都市立松ヶ崎小学校の来栖朝子教諭に感謝する．また，被験者になっていただいた方々に謝意を表する．本研究の一部は京都高度技術研究所の平成 12 年および平成 15 年学生ベンチャー奨励金制度によるものである．

参 考 文 献

1) Kim, J.-Y. and Belkin, N.J.: Categories of music description and search terms and phrases used by non-music experts, *Proc. 3rd International Conference on Music Information Retrieval* (*ISMIR2002*), Paris, France, pp.209–214 (2002).
2) 木本晴夫：感性語による画像検索とその精度評価，情報処理学会論文誌，Vol.40, No.3, pp.886–898 (1999).
3) 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.40, No.5, pp.2356–2366 (1999)
4) 栗田多喜夫，加藤俊一，福田郁美，坂倉あゆみ：印象語による画像データベースの検索，情報処理会論文誌，Vol.33, No.11, pp.1373–1383 (1992).
5) 清木 康，金子昌史，北川高嗣：意味の数学モデルによる画像データベース探索方式とその学習機構，信学論，Vol.J79-D-II, No.4, pp.509–519 (1996).
6) 井上光平，浦浜喜一：混合密度回帰に基づく感性検索，信学論，Vol.J83D-II, No.4, pp.1192–1194 (2000).
7) Takatera, M., Furukawa, T., Shimizu, Y., Kamijo, M., Hosoya, S., Morisaki, T. and Ohtake, A.: Apparel Products Search System Considering Individual Kansei Evaluation, *KANSEI Engineering International*, Vol.1, No.2, pp.1–8 (2000).
8) 角 薫，角 康之，間瀬健二，中須賀真一，堀浩一：個人の概念空間を利用した興味の推定による情報提供，信学論，Vol.J82-D-II, No.10, pp.1634–1644 (1999).
9) 國分三輝，倉橋哲郎，古西浩之，向江秀之，井口弘和，川澄未来子：個人感性情報を用いた住空間設計支援システム，ヒューマンインタフェース学会論文誌，Vol.3, No.4, pp.239–248 (2001).
10) 北川高嗣，中西崇文，清木 康：楽曲メディアデータを対象としたメタデータ自動抽出方式の実現とその意味的楽曲検索への応用，信学論，Vol.J85D-I, No.6, pp.512–526 (2002).
11) 池添 剛，梶川嘉延，野村康雄：音楽感性空間を用いた感性語による音楽データベース検索システム，情報処理学会論文誌，Vol.42, No.12, pp.3201–3212 (2001).
12) 谷口高士：音楽と感情，北大路書房 (1998).
13) 岩下豊彦：SD 法によるイメージの測定，川島書店 (1983).
14) 坂本 崇，梶川嘉延，野村康雄：音楽感性空間における非線形判別分析をもちいた曲印象別グループの分割，情報処理学会論文誌，Vol.40, No.4, pp.1901–1909 (1999).
15) 小林重順：カラーリスト—色彩心理ハンドブック，講談社 (1997).
16) Sugihara, T., Morimoto, K. and Kurokawa, T.: An Improved Kansei-Based Music Retrieval System with a New Distance in a Kansei Space, *Proc. 13th IEEE International Workshop on Robot and Human Interactive Communication* (*RO-MAN2004*), Kurashiki, Okayama, pp.141–146 (2004).

(平成 16 年 10 月 14 日受付)
(平成 17 年 5 月 9 日採録)

杉原　太郎

1998 年徳山高専機械電気工学科卒業．2000 年同高専専攻科機械制御工学専攻修了．2002 年京都工芸繊維大学大学院工芸科学研究科博士前期課程修了．2005 年同研究科博士後期課程修了．博士（工学）．同年北陸先端科学技術大学院大学知識科学研究科助手．感性検索システムの開発および感性情報処理に関する研究に従事．ヒューマンインタフェース学会，日本感性工学会，日本音楽知覚認知学会，ACM 各会員．

森本　一成（正会員）

1978 年京都工芸繊維大学工業短期大学部技官教務職員．1994 年同大学工芸学部講師．1998 年同助教授．同年同大学大学院工芸科学研究科助教授．現在に至る．工学博士．ヒューマンインタフェースの設計と評価に関する研究に従事．電子情報通信学会，照明学会，日本人間工学会，日本デザイン学会，画像電子学会，映像情報メディア学会等の会員．

黒川　隆夫（正会員）

1966 年大阪大学基礎工学部制御工学科卒業．1971 年同大学大学院基礎工学研究科博士課程修了．工学博士．同年同大学基礎工学部助手．京都工芸繊維大学工芸学部助教授，教授等を経て，1998 年同大学大学院工芸科学研究科教授．この間，適応制御，生体工学等の研究に従事．現在は認知工学，メディア工学，ヒューマンインタフェース，感性工学が主テーマ．著書『ノンバーバルインタフェース』（オーム社),『自己の表現』（岩波書店）等．電子情報通信学会，ヒューマンインタフェース学会，日本人間工学会，IEEE，ACM 等の会員．