

## 風景画像とサンプル楽曲を用いた 環境状況コンテキスト対応型音楽推薦システムの実現

桐本 篤<sup>†</sup> 佐々木 史織<sup>‡</sup> 清木 康<sup>†</sup>

<sup>†</sup> 慶應義塾大学環境情報学部 〒252-0011 神奈川県藤沢市遠藤 5322

<sup>‡</sup> 慶應義塾大学政策・メディア研究科 〒252-0011 神奈川県藤沢市遠藤 5322

E-mail: <sup>†</sup> {t06305ak, kiyoki} @sfc.keio.ac.jp, <sup>‡</sup> {sashiori} @mdbf.sfc.keio.ac.jp

**あらまし** 本稿では、ユーザの入力した風景画像の色彩情報から環境状況コンテキストを判定し、同時に、ユーザの選択したサンプル楽曲からユーザの嗜好・感性を抽出し、音楽感性属性を媒介とした環境状況コンテキストと音楽感性データの相関量計量により、コンテキストに相関が高く、個人の感性に合致した音楽データを推薦・提示するシステムの実現方式について示す。本稿では、空の画像を実風景として、それに相関の高い環境状況コンテキスト（ここでは天候）を判定し、その天候に対応する感性データの抽出を行い、その感性データと相関が高く、また個人の音楽嗜好に合致する音楽データの自動抽出を行う実験を示し、本方式の実現可能性を明らかにする。

**キーワード** 情報推薦, マルチメディア, パーソナライゼーション, 意味的連想検索

## An Environmental Context-Responsive Music Recommendation System with Scenic Images and Sample Music

Atsushi KIRIMOTO<sup>†</sup>, Shiori SASAKI<sup>‡</sup> and Yasushi KIYOKI<sup>†</sup>

<sup>†</sup> Faculty of Environmental Information, Keio University, Endo 5322, Fujisawa, Kanagawa, JAPAN

<sup>‡</sup> Graduate School of Media and Governance, Keio University, Endo 5322, Fujisawa, Kanagawa, JAPAN

E-mail: <sup>†</sup> {t06305ak, kiyoki} @sfc.keio.ac.jp, <sup>‡</sup> sashiori@mdbf.sfc.keio.ac.jp

**Abstract** This paper presents an implementation method of an environmental context-responsive music recommendation system. First, a user's environmental context is identified by a scenic image input by the user in this system. Second, a user's preference of music is extracted by sample music data selected by the user. Third, the correlations between environmental context and music data are calculated through the musical "kansei", and the highly-correlated music data to the mood of user are selected and recommended. To examine the feasibility of this system, we have performed several experiments by setting images of sky as environmental data, and two types of user as model users with different musical preferences.

**Keyword** Information Recommendation, Multimedia, Personalization, Semantic Associative Search

### 1. はじめに

#### 1.1. 問題の所在

近年の携帯音楽プレイヤーの大容量小型化と、音声圧縮技術の高度化はめざましく、音質も携帯性も非常に向上し、大量の音源データを持ち運ぶことのできるモデルが存在している。そのような携帯音楽プレイヤー利用において、大量の音楽データの中から、自然環境などの状況に応じて楽曲を動的に選択し、それらを活用することのできる利用環境を実現することは、音楽データの新たな利用、展開の場を提供するという視点において重要である。

一方、ユーザの嗜好というものは現実には非常に流動的である。真夏の夜の、雪の降った明け方に聞きたい楽曲が必ずしも一致しないことは明らかであり、例えば晴れた日には明るい曲を、雨の日には静かな曲を、昼には活発な印象を与える曲を、夜には暗い曲を求めるといったように、一人のユーザであっても時間や天候の影響によって気分、あるいは嗜好が変化していくと考えられる。しかし従来からの嗜好にもとづく楽曲推薦システムは、こうしたユーザの日々の時空間に起因する要求の変化に対して十分に対応しているとは言えない。ユーザの流動的な嗜好の変化をとらえることができれば、現在普及している携帯電話や携帯音楽プレイヤーなどのデバイスをよりスマートに活用できる

と考えられる。(図1)

#### 1.2. 関連研究

情報推薦の分野においては、協調フィルタリングや推薦メディアの内容によるフィルタリングに関する研究がある[1][2]。特に、音楽推薦システムについては、協調フィルタリングによる楽曲プレイリスト生成方式[3]、協調フィルタリングとユーザプロフィール編集を組み合わせた楽曲推薦システム[4]、内容に基づくフィルタリングを用いた音楽推薦方式[5]、アーティストなどの情報を利用したプレイリスト生成方式[6]などが提案されており、また、内容に基づくフィルタリングと協調フィルタリングの連動による感性情報を用いた楽曲推薦システム[7]も研究されている。

本研究は、色彩情報と感性情報の特徴としてユーザの環境状況と推薦メディアデータの関連づけを行い、これをユーザプロフィールと内容に基づくフィルタリングとして用いる研究として位置づけることができる。

#### 1.3. 研究の目的と特徴

本研究の目的は、静的なユーザの嗜好ではなく、時空間に依存した要求を情報推薦の基準とすることで、生きているユーザに対して従来よりもきめ細かくアプローチするシステムの実現にある。

また、本研究の特徴は、ユーザを取りまく環境をセンシングするうえで、ユーザが携帯カメラ等によって

取得した風景画像の色彩ヒストグラムから感性情報を抽出し、また、その感性情報を媒介に、選択された楽曲群を個人の感性嗜好データとして蓄積することにより、ユーザの環境状況と感性に合致した情報を推薦するシステムを実現する点にある。

環境をセンシングする、という研究はこれまでも様々な形で行われてきたが、色彩ヒストグラムの比較を用いることによって非接触かつデータの質にとらわれず環境をセンシングすることが可能となる。ひいては、環境状況コンテキストに対する人間の感情的レスポンスを「感性情報」としてあらかじめ設定しておくことで、人間の感情の推移や嗜好の変化を間接的にセンシングすることも可能である。また、データベースの蓄積が容易であることから、データの内容を変更し、相当量のデータベースを構築することさえできれば、オープンワールドに非常に近いレベルでの環境のセンシングが可能となる。

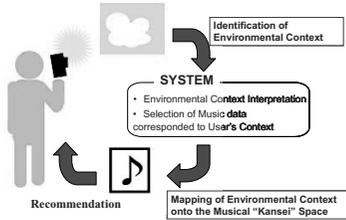


図1 基本アイデア

## 2. 基本方式

本システムは、携帯カメラ等によって取得した風景画像を色彩ヒストグラムとして分析し、事前に用意しておいた任意の数のパターンの風景画像（以下、サンプル画像データと呼ぶ）から抽出した色彩データベースとの相関量を計量し、相関量の高い風景画像の状況を現在の環境状況コンテキストとして認識する。

本方式は、大きく分けて次の2ステップにより構成される。

ステップ1: いくつかの環境状況データが設定され、これら環境情報に基づいて選択された楽曲群を元に、あらかじめ環境状況コンテキストには感性情報が付与される。この環境状況コンテキストに付与された感性情報は、ユーザごとに生成されることにより、ユーザの感性と嗜好に沿った情報を可能とする。

ステップ2: 認識された環境状況コンテキストに付与された感性情報と楽曲データの感性情報間での相関量計量（ここでは内積計算を用いる）を行い、相関量の高い順にあらかじめ指定した数の音源情報を取得し、それをもとにプレイリストを作成する。

### 2.1. システム構成

本システムは、次の2機能によって構成される。

- (機能1) 環境状況判定機能
- (機能2) 感性情報マッピング機能

入出力データと各機能、および、データの流れを図2のシステム構成図に示し、以下の節にて各機能の詳細について述べる。

#### 2.1.1. 環境状況判定機能

##### STEP0: 色彩分布情報抽出

入力された判定対象画像ファイルについて、画像の各画素の色をシステムで設定された基本色にマッピングし、画像の特徴をしめすカラーヒストグラムを作成する。ここでカラーヒストグラムとは、画像中の各基本色の正規化された画素数を要素として有するベクトルとして表現される。

##### STEP1: 色彩-環境状況データベースの作成

予め用意された、 $p$  個の環境状況 ( $I_1, I_2, \dots, I_p$ ) を代

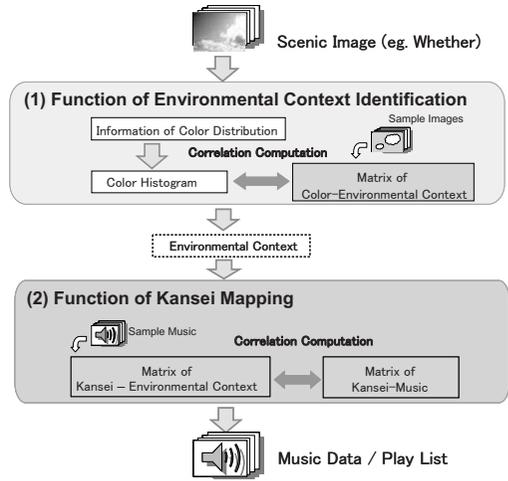


図2 システム構成図

表する  $n$  個のサンプル画像について、STEP0の手順でカラーヒストグラムを作成し、マトリクスの形でデータベース化する（以下、「色彩-環境状況マトリクス」と呼ぶ）。色彩-環境状況マトリクスとは、色彩と環境状況との関連付けを数値として示すものであり、各環境状況は、 $n$  個のサンプル画像データによって色彩との相関の強さによって重みづけられている。

色彩-環境状況判定に用いる色は、システムとして選択した  $m$  個の基本色である。この基本色は、環境状況を表現するのに十分な数が必要である。色彩-環境状況マトリクスにおける環境状況は、必要に応じて追加や削除が可能である。

この、各環境状況を表現するサンプル画像の色彩分布パターンに現れる、 $m$  個の基本色の頻度を数値化したものを色彩-環境状況ベクトル  $s = \{s_1, s_2, \dots, s_m\}$  ( $I$  は環境状況識別子) とする。ここで、色彩-環境状況ベクトルを行ベクトルとする  $n$  行  $m$  列のマトリクスを色彩-環境状況マトリクス  $C$  とする(図3)。

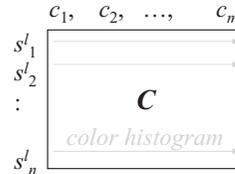


図3 色彩-環境状況マトリクス  $C$

##### STEP2: 判定対象画像のカラーヒストグラム作成

判定対象画像の特徴量として、STEP1の手順において作成した色彩-環境状況マトリクス  $C$  の  $m$  個の基本色についてカラーヒストグラムを作成する。ここでは、各画像  $d$  における各色  $c_j$  ( $j=1,2,\dots,m$ ) の画素数を求め、各画像の全画素数で正規化し、面積比として算出したものを用いる。このカラーヒストグラムを画像の特徴ベクトル  $h = \{h_1, h_2, \dots, h_m\}$  とする。

##### STEP3: 色-環境状況マトリクスを用いた相関量計量

判定対象画像の特徴を示すカラーヒストグラムについて、色彩-環境状況マトリクス  $C$  を用いて、ユーザの環境状況の判定を行う。環境状況判定時は、画像から直接作成された、色の使用面積からなるカラーヒストグラムだけではなく、STEP2において作成された色

彩-環境状況マトリクス, すなわち, サンプル画像群のカラーヒストグラムから色が環境状況に与える重みを付けられた重み付けカラーヒストグラムを使用する.

色彩-環境状況マトリクス  $C$  の行ベクトル  $s = \{s_1, s_2, \dots, s_m\}$  は, 環境状況に対する色彩分布の傾向を表現したもになっている. 判定対象画像ベクトル  $h$  は色彩-環境状況ベクトル  $s$  と同一の特徴空間にマップすることができるため, ベクトル空間上で対象画像と色彩-環境状況ベクトルの相関量を計量することが可能となる. 判定対象画像  $I$  と環境状況  $s_k$  の相関量  $I_k$  は, 画像の特徴ベクトル  $h$  と色彩-環境状況ベクトル  $s_k = \{s_{k1}, s_{k2}, \dots, s_{km}\}$  を用いて次のように計算できる. 相関量は, 印象語の重みを正規化するためにコサイン尺度を用いる.

$$\cos(I, s_k) = \frac{\sum_{j=1}^m h_j \cdot s_{kj}}{\sqrt{\sum_{j=1}^m h_j^2} \cdot \sqrt{\sum_{j=1}^m s_{kj}^2}}$$

$h_j$ : 画像  $I$  の色  $c$  へのヒストグラム要素  
 $s_{kj}$ : 環境状況サンプルの色  $c$  への重み

ここで, 判定対象画像と相関量が高い環境状況サンプル画像が「環境状況コンテキスト」として抽出される.

### 2.1.2. 感性情報マッピング機能

#### STEP1: 感性-音楽データベースの作成

予め用意された,  $o$  個の楽曲データ ( $M_1, M_2, \dots, M_o$ ) について, 感性を表す  $q$  個の感性語 ( $e_1, e_2, \dots, e_q$ ) と楽曲データの関連をマトリクスの形でデータベース化する (以下, 「感性-音楽マトリクス」と呼ぶ). 感性-音楽マトリクスは, 感性と音楽データとの関連付けを数値として示すものであり, 各音楽データは  $q$  個の感性語 ( $e_1, e_2, \dots, e_q$ ) によって感性との相関の強さが数値として表現されている. この, 各音楽データを特徴づける  $q$  個の感性語との相関を数値化したものを感性-音楽ベクトル  $M_j = \{e_{j1}, e_{j2}, \dots, e_{jq}\}$  とする. ここで, 感性-音楽ベクトルを行ベクトルとする  $o$  行  $q$  列のマトリクスを感性-音楽マトリクス  $M$  とする (図 4).

#### STEP2: 感性-環境状況データベースの作成

予め用意された  $p$  個の環境状況 ( $l_1, l_2, \dots, l_p$ ) について, ユーザによって数曲の楽曲データが, 状況に合致したサンプル楽曲として選択される. このサンプル楽曲から抽出された, 感性を表す  $q$  個の感性語 ( $e_1, e_2, \dots, e_q$ ) と環境状況との関連をマトリクスの形でデータベース化する (以下, 「感性-環境状況マトリクス」と呼ぶ).

各環境状況を特徴づける  $q$  個の感性語との相関を数値化したものを感性-環境状況ベクトル  $l_k = \{e_{k1}, e_{k2}, \dots, e_{kq}\}$  とする. ここで, 感性-環境状況ベクトルを行ベクトルとする  $p$  行  $q$  列のマトリクスを感性-環境状況マトリクス  $L$  とする (図 4). 感性-環境状況マトリクス  $L$  は, 感性と環境状況との関連付けを数値として示すものであり, 各環境状況は STEP1 で用いられる感性語によって感性との相関の強さが数値として表現されている.

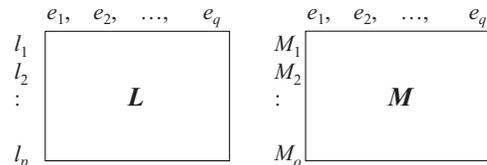


図 4 感性-環境状況マトリクス  $L$  および, 感性-音楽マトリクス  $M$

### STEP3: 感性特徴による環境状況と音楽データの相関量計量

感性-環境状況ベクトル  $I_x$  は感性-音楽ベクトル  $M_j$  と同一の特徴空間にマップすることができ, 2.1.1 で同定された環境状況情報と各音楽データはベクトル空間上で相関量を計量することが可能となる. 判定対象環境状況  $I_x$  と音楽データ  $M_j$  の相関量は, 次のように内積  $\sigma$  によって計算することとなる.

$$\sigma(I_x, M_j) = \sum_{i=1}^q e_{xi} \cdot e_{ji}$$

### 3. 実現方式

ここでは, 環境状況コンテキストを「天候」に設定し, 空の画像を実風景として, それに相関の高い環境状況 (天候) を判定し, その天候に対応する感性データの抽出を行い, その感性データと相関の高い音楽データの自動抽出を行うシステムの実現方式について述べる.

#### 3.1. 環境状況判定機能

天候に関して, 晴れ (fine)・曇り (cloudy)・夕焼け (sunset)・夜 (night)・朝焼け (daybreak)・雪 (snow)・雨 (rain) などのパターンを複数設定する. この複数パターンの各天候につき数十件ずつ, Google Image Search [10]などの画像検索エンジンや Flickr [11]などの画像共有アーカイブを用いてサンプル画像を取得し, そのヒストグラムを用いてマトリクスの形で色彩-環境状況データベースを作成する.

色彩-環境状況マトリクス生成のための基本色としては, カラーイメージスケール [8]による有彩色 120 色, および, 無彩色 10 色の計 130 色の色を用いる. サンプル画像データおよび判定対象画像データのヒストグラム生成は, 画像の各画素の RGB 値について HSV 値に変換し, カラーイメージスケールの代表色 130 色の RGB 値に対して, HSV 空間の円錐モデルで最も近い色を選択して, 画像全体の 130 色の占める割合を算出した [9]. また, 環境状況コンテキストを読み込むための判定対象画像データは, 想定ユーザが撮影したものを利用する.

#### 3.2. 感性情報マッピング機能

第一に, 推薦対象の音楽メディアについて, Hevner [12]が設定した音楽の印象を表現するための 8 印象語群 C1~C8 (図 5) を感性語として用いて, 感性-音楽データベースを構築する. 楽曲 Midi ファイルに対して文献 [13]に述べられている感性情報自動抽出手法を用いて解析する. それらの楽曲の Midi を文献 [13]に述べられている感性情報自動抽出手法を用いて解析する. この解析結果をもとに, C1 から C8 までの重み付き印象語と楽曲データとの関連が感性-音楽マトリクスとして生成される.

<p><b>C8</b>  emphatic 強調された  exalting 誇り高い  majestic 荘厳な  marial 美しい, 荘厳な  ponderous 大きく重い, 重々しい  robust 強壯な  vigorous 精力旺盛な, 活気のある</p>	<p><b>C7</b>  agitated 動揺した, 興奮した  dramatic ドラマチック, 劇的な  exciting 手に汗を流す, 面白い  enflamed 炎上した  impetuous 猛烈な  passionate 熱意な, 情熱的な  restless 落ち着かない, 眠れない  sensational 騒がせる, 驚異的な  soaring 上がり止まり</p>	<p><b>C6</b>  bright 明るい  cheerful 楽しい  gay 楽しい  happy 楽しい  joyous 楽しい  merry 楽しい</p>	<p><b>C5</b>  delicate 繊細な  fanciful ロマンチックな  grandful 豪華な  humorous 滑稽な, ユーモラスな  light 軽快な  playful 遊び心のある  quaint 風変わりでおもしろい, 古風な  spritly 活発な, 元氣な  whimsical 気まぐれな, 変な</p>	<p><b>C4</b>  calm のんぼり, 悠静  leisurely ゆっくりした, 急がない  lyrical 韻律的, 音楽的な  quiet 静かな  satisfying 満足な, 十分な  serene 穏やかな, 落ち着いた  soothing 落ち着かせる, 心地よい  tranquil 静かな, 平穏な</p>
<p><b>C1</b>  awe 恐れ, 畏怖  inspiring 鼓舞する, 意気をもたせる  digitized 感動的な, 高貴な  lofty 高貴な, 堂々とした  sacred 神聖な  serious 真面目な  sober 謙遜な  solemn 厳格な</p>	<p><b>C2</b>  dark 暗い, 薄暗い  depressing 鬱陶しい  dreadful 恐ろしい  frustrated 挫折した, げげられた  gloomy (薄暗い), 憂鬱な  heavy 重苦しい, 物陰しい  melancholy 陰気, 陰鬱な  mourning 悲しみに耽んだ</p>	<p><b>C3</b>  dreamy 夢の多い, 夢みりするような  longing 憧れ, 渇望  plaintive 悲しげな, 哀愁に満ちた  sentimental 感傷的な, 情にふける  tender 柔らかな, 優れやうしい  yearning 憧れ, 思慕  yielding 寛容な, しなやかな</p>		

図 5 Hevener [1937]による印象語群

第二に、3.1において設定した複数パターンの環境状況(天候)について、Hevnerの8印象語群C1~C8(図5)を感性語として用いて感性-環境状況データベースを構築する。想定ユーザに対して、3.1において設定した複数パターンの環境状況に合致する楽曲を複数挙げてもらい、それらの楽曲のMidiを文献[13]に述べられている感性情報自動抽出手法を用いて解析する。この解析結果をもとに、C1からC8までのそれぞれの重み付き印象語における重みの最大値を採用し、合成することで、感性-環境状況ベクトルとする。これにより、環境状況と印象語群(C1~C8)の関連性を数値で記述し、マトリクスの形で感性-環境状況データベースを自動的に構築することができる。

#### 4. 実験

実験環境として、3.1において記述した方法で、晴れ(fine)・曇り(cloudy)・夕焼け(sunset)・夜(night)・朝焼け(daybreak)・雪(snow)・雨(rain)の7パターンの天候について、サンプル画像を10件ずつ、Google Image Search[10]およびFlickr[11]を用いて計70件取得し、そのヒストグラムを用いて色彩-環境状況データベースを作成した。

環境状況判定機能の検証である実験1では、判定対象画像データとして想定ユーザが携帯カメラで撮影した画像を利用して。また、感性情報マッピング機能の検証である実験2では、推薦対象音楽データとしてBeatlesの楽曲Midiを210曲分使用した。最終的に生成されたプレイリストの被験者による評価、および、感性-環境状況データベース作成の際のサンプル音源には、Midi音源ではなく、対応する楽曲のCD音源を使用した。アウトプットされる楽曲リストは各環境状況に対して20曲分とした。

##### 4.1. 実験1: 環境状況判定機能の検証

環境状況(天候)判定の基準として、次の二つの基準を設定し目視での結果予想との比較を行った。

(判定基準1) 最も相関量の高いサンプル画像の環境状況タグを環境状況コンテキストとして判定する。

(判定基準2) 環境状況タグごとにサンプル画像の相関量を合計し、最も高い値をもつ環境状況タグを環境状況コンテキストとして判定する。

結果の一部を表1に示す。

Input Image (Query)	Expected Result	Rank: Correlation Value	Criterion 1	Criterion 2
Q1 	fine	fine4: 0.9556 fine2: 0.9555 fine7: 0.9461 fine8: 0.9371 fine6: 0.5991 daybreak5: 0.55105 rain1: 0.20657 snow7: 0.2822 rain0: 0.23994 sne wd: 0.22204 sunset8: 0.85086 sunset5: 0.53664 night3: 0.51868 night2: 0.51868 night9: 0.51624 night4: 0.51481 night0: 0.51481 sunset9: 0.48606 daybreak1: 0.48606 night8: 0.48334	fine	fine: 2.51906 daybreak: 0.55105 rain: 0.54651 snow: 0.50524
Q2 	sunset	sunset8: 0.85086 sunset5: 0.53664 night3: 0.51868 night2: 0.51868 night9: 0.51624 night4: 0.51481 night0: 0.51481 sunset9: 0.48606 daybreak1: 0.48606 night8: 0.48334	sunset	night3: 0.6656 sunset: 1.68356 daybreak: 0.48606
Q3 	night	night1: 0.99812 night8: 0.98683 night4: 0.93196 night0: 0.93196 night3: 0.80811 night2: 0.80811 sunset7: 0.79733 night9: 0.78537 night5: 0.78391 daybreak5: 0.78221	night	night7: 0.0417 sunset: 0.79733 daybreak: 0.75221
Q4 	daybreak	fine1: 0.73921 cloudy8: 0.69889 cloudy1: 0.68103 daybreak6: 0.67872 cloudy9: 0.46419 snow5: 0.62822 rain7: 0.48109 rain0: 0.47978 snow9: 0.47638 rain6: 0.46209	fine	cloudy2: 0.02411 snow: 1.56849 rain: 0.96095 fine: 0.73921 daybreak: 0.67872

表1 実験1: 環境状況判定の結果

判定対象画像 Q1, Q3, に関しては、判定基準1, 判

定基準2のいずれによっても目視予想と合致した判定結果が得られた。Q2, については、判定基準1による結果(sunset/night)の方が目視予想と合致しているが、これは判定基準2において樹木の影の部分の黒にデータベース内の黒い画像、つまり夜の画像が多く反応を示したためである。いずれにしても妥当な結果といえる。一方、Q4については基準1による結果(fine)と基準の結果(cloudy/snow)と、いずれも目視予想から外れた判定結果が得られたが、Q4については太陽の光が白く映っている部分や上部の青空の部分の白色が影響していると考えられる。しかし、Q4のような誤判定については、サンプル画像のデータ量を増やし、例外的な画像をデータベースに追加することで、十分に精度の向上が可能と考えられる。

##### 4.2. 実験2: 感性情報マッピング機能の検証

被験者A, Bの二種類の被験者を設定し、3.2において記述したプロセスにより、実験1で用いた7パターンの天候とそのサンプル画像を示しながら、これら天候に合致するBeatlesの楽曲を複数挙げてもらい、それらの楽曲Midiから自動抽出された感性情報(重み付き印象語群C1~C8)の重みを合成することで感性-環境状況ベクトルを作成した。

合成ベクトルの作成にあたり、重みの(1)最大値を採用、(2)最小値を採用、(3)平均値を採用、(4)閾値以上の最大値を採用、の4種類の計算方法により、感性-環境状況の合成ベクトルを生成した。最大値を採用する場合は、環境状況に僅かでも対応する楽曲を網羅的に検索することに適していると前提に立つ。一方、最小値を採用する場合は、逆にわずかな感性情報をノイズとして扱い、最低限のデータを利用することによってノイズを除去することに適していると期待される。平均値では、同一の環境状況で挙げられたサンプルであれば、同一の傾向を持っているであろうという仮説のもと、全体的な傾向の抽出を行う。閾値以上の最大値を採用する場合は、部分的に最小値を採用することでノイズを削減し、そうでない部分には最大値を採用することで、挙げられたサンプル楽曲群の感性情報を強調して抽出する。

以上の4つの計算方法により、サンプル楽曲から抽出された感性情報を元に感性-環境状況ベクトルを自動生成した例を、表2および表3に示す。

表2: 被験者Aが「晴れ」に合う楽曲として選択したサンプル楽曲と抽出された感性情報、および、その値を用いて生成された感性-環境状況ベクトル

Impression	C1	C2	C3	C4	C5	C6	C7	C8
Music ID	0.061769	0	0	0	0	0.135699	0.481781	0.320751
can't buy	0.061769	0	0	0	0	0.135699	0.481781	0.320751
ob-la-di	0.677202	0	0	0.177542	0	0	0	0.145256
yellow	0.275276	0	0	0.044649	0.310325	0.18391	0.054892	0.130948

Impression	C1	C2	C3	C4	C5	C6	C7	C8
kansel vector	0.677202	0	0	0.177542	0.310325	0.135699	0.481781	0.320751
(1) by max	0.677202	0	0	0.177542	0.310325	0.135699	0.481781	0.320751
(2) by min	0.275276	0	0	0	0	0	0	0.130948
(3) by average	0.338062	0	0	0.074063	0.103442	0.106536	0.178891	0.198985
(4) by threshold	0.677202	0	0	0	0.310325	0	0.481781	0.320751

表3: 被験者Aが「夕暮れ」に合う楽曲として選択したサンプル楽曲と、生成された感性-環境状況ベクトル

Impression	C1	C2	C3	C4	C5	C6	C7	C8
Music ID	0.51508	0	0	0.098752	0.166798	0.024956	0	0.194416
something	0.51508	0	0	0.098752	0.166798	0.024956	0	0.194416
let it be	0.269779	0	0.207774	0.430152	0.092295	0	0	0
imstire	0	0	0.343772	0.44735	0.208878	0	0	0

Impression	C1	C2	C3	C4	C5	C6	C7	C8
kansel vector	0.51508	0	0.343772	0.44735	0.208878	0.024956	0	0.194416
(1) by max	0.51508	0	0.343772	0.44735	0.208878	0.024956	0	0.194416
(2) by min	0	0	0	0.098752	0.092295	0	0	0
(3) by average	0.26162	0	0.183849	0.325418	0.155989	0.008319	0	0.064805
(4) by threshold	0.51508	0	0.343772	0.44735	0.092295	0	0	0

表 4：被験者 A のサンプルから生成された 4 種類の感性-環境状況ベクトルによる  
楽曲検索結果の比較例（環境状況「曇り」と「晴れ」の場合）

[cloudy]				[sunset]				[average]				[threshold]			
ranking	track	correlation	points	ranking	track	correlation	points	ranking	track	correlation	points	ranking	track	correlation	points
[1]	goodnite:	0.42849459		2	ififeel:	0.082295395		2	blkbird:	0.249972555		0	blkbird:	0.423013727	0
[2]	blkbird:	0.423013727		2	blujaywy:	0.081372567		2	eleonor:	0.248824663		2	herether:	0.422055775	2
[3]	herether:	0.422186689		2	tomorrow:	0.081208758		1	idntwant:	0.245776337		0	goodnite:	0.421265324	2
[4]	eleonor:	0.419428951		2	golden:	0.080932308		2	yesterday:	0.244277275		2	eleonor:	0.419428951	2
[5]	lovrita:	0.417849827		1	yesterday:	0.080758328		2	goodnite:	0.244199529		2	idntwant:	0.416598019	0
[6]	idntwant:	0.416598019		0	flying:	0.078702911		2	ivegot:	0.244081743		2	shecame:	0.413489196	0
[7]	shecame:	0.413948729		1	ivegot:	0.076711961		2	golden:	0.242650421		2	youbegot:	0.411830937	2
[8]	youbegot:	0.412864439		2	iwantto:	0.076617614		0	ififeel:	0.241928381		2	follosn:	0.411139835	2
[9]	follosn:	0.411139835		2	altheip:	0.076356513		2	blujaywy:	0.241132248		2	ivegot:	0.409314197	2
[10]	hellybye:	0.410461006		2	bemrkite:	0.076095656		2	iffollosn:	0.24105421		2	hellybye:	0.409019483	2
				0.7				0.8				0.8			

[sunset]				[average]				[threshold]							
ranking	track	correlation	points	ranking	track	correlation	points	ranking	track	correlation	points				
[1]	hellybye:	0.453012108		0	hellybye:	0.318228534		2	hellybye:	0.453012108	2				
[2]	onlyanor:	0.440592106		2	polypam:	0.316448822		0	onlyanor:	0.436865229	2				
[3]	polypam:	0.436816614		0	NORWOOD:	0.205363822		2	polypam:	0.433905753	0				
[4]	sunking:	0.430517279		2	letitbe:	0.2036058		2	youbegot:	0.423586923	2				
[5]	youbegot:	0.429724465		2	imwalrus:	0.203307821		2	sunking:	0.417437641	2				
[6]	longlong:	0.424275957		2	dntletme:	0.201111266		1	sunking:	0.410344537	2				
[7]	rockeyrac:	0.416627544		2	eightday:	0.2004686		0	letitbe:	0.392985716	2				
[8]	hejude:	0.414739735		2	wheniget:	0.200335691		2	NORWOOD:	0.391292313	2				
[9]	dearprud:	0.409556407		2	lowyouto:	0.200356555		2	sexsyad:	0.390893768	2				
[10]	crybacry:	0.406337167		2	hellybye:	0.199643771		2	rockeyrac:	0.389671676	2				
				0.9				0.75				0.9			

表 5：被験者 B のサンプルから生成された 4 種類の感性-環境状況ベクトルによる  
楽曲検索結果の比較例（環境状況「曇り」と「晴れ」の場合）

[cloudy]				[sunset]				[average]				[threshold]			
ranking	track	correlation	points	ranking	track	correlation	points	ranking	track	correlation	points	ranking	track	correlation	points
[1]	glassoni:	0.413378707		2	glassoni:	0.278384971		2	glassoni:	0.347291039		2	glassoni:	0.412718655	2
[2]	IFINEED:	0.40419032		1	babytsu:	0.263587338		1	babytsu:	0.332590867		1	babytsu:	0.392197488	1
[3]	ticketrd:	0.4038581		2	IFINEED:	0.259685135		1	IFINEED:	0.331892667		1	yourmoth:	0.38679176	2
[4]	devilin:	0.402489945		2	dizzlizz:	0.259339134		1	ticketrd:	0.328340696		2	IFINEED:	0.384982592	1
[5]	babytsu:	0.400878527		1	yourmoth:	0.258881185		2	YOWNTSEE:	0.326646884		2	dizzlizz:	0.383813845	0
[6]	DRIMYCAR:	0.40052642		1	YOWNTSEE:	0.256055328		2	itsonly:	0.326152392		1	YOWNTSEE:	0.381333777	2
[7]	itsonly:	0.399848005		1	ticketrd:	0.254094434		2	DRIMYCAR:	0.326091686		1	tsonly:	0.379090416	1
[8]	savoytr:	0.39903956		1	itsonly:	0.253775757		1	devilin:	0.325921245		2	ticketrd:	0.377482851	2
[9]	YOWNTSEE:	0.397583387		2	DRIMYCAR:	0.252890833		1	yourmoth:	0.325605955		2	wantell:	0.375247325	2
[10]	badboy:	0.397390776		1	devilin:	0.251112825		2	dizzlizz:	0.324520932		0	DRIMYCAR:	0.373565955	1
				0.7				0.75				0.7			

[sunset]				[average]				[threshold]							
ranking	track	correlation	points	ranking	track	correlation	points	ranking	track	correlation	points				
[1]	hellybye:	0.468013685		1	polypam:	0.277231534		0	hellybye:	0.468013685	2				
[2]	polypam:	0.455064015		1	hellybye:	0.27582612		2	polypam:	0.451387126	0				
[3]	youbegot:	0.45198314		2	younever:	0.077976499		2	onlyanor:	0.450661615	1				
[4]	onlyanor:	0.450661615		1	revolu l:	0.07431926		1	youbegot:	0.444214341	2				
[5]	sunking:	0.431296759		1	WHATGOES:	0.070749928		0	shecame:	0.431296759	1				
[6]	longlong:	0.423273628		2	comatoge:	0.070199274		2	letitbe:	0.423273628	2				
[7]	letitbe:	0.422091776		2	wantyouy:	0.068816417		1	herether:	0.411331795	2				
[8]	shecame:	0.42164226		1	honeydnt:	0.067249612		0	sunking:	0.255226902	2				
[9]	eleonor:	0.407284839		1	meanmr:	0.066898066		1	longlong:	0.254915953	2				
[10]	yesterday:	0.40647385		2	twontbe:	0.066422936		0	imwalrus:	0.25380274	1				
				0.7				0.45				0.7			

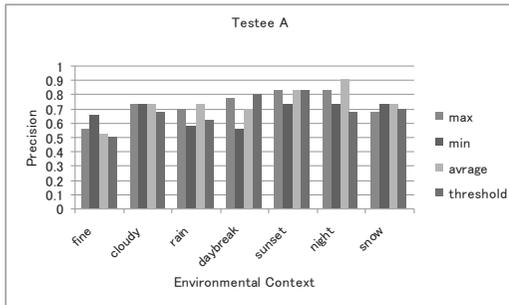


図 6：各環境状況に対する楽曲検索結果適合率  
(被験者 A の場合)

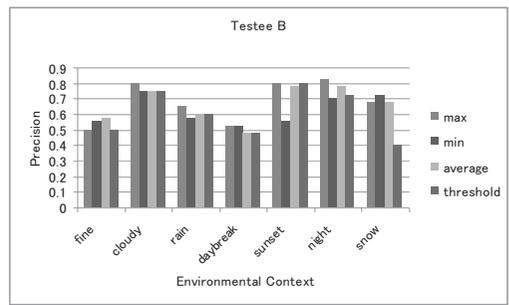


図 7：各環境状況に対する楽曲検索結果適合率  
(被験者 B の場合)

表2は、被験者Aが「晴れ」に合う楽曲として選択したサンプル楽曲3つと、それらから抽出されたC1~C8の感性情報、および、その値を用いた4つの方法による合成ベクトルを示している。同様に、表3は、被験者Aが「夕暮れ」に合致するとして選択したサンプル楽曲3と感性情報、合成ベクトルの例である。表2と表3を比較すると、表2では「堂々とした、厳粛な、神聖な」といったC1や「ドラマチックな、活気づいた」といったC7の要素に値の大きいベクトルが生成されているのに対し、表3では「哀愁に満ちた、感傷的な」といったC3や「ゆっくりした、物静かな」のC4の要素に値の大きいベクトルが生成されていることが分かる。

この感性・環境状況ベクトルを用いて相関量計算を行った結果をもとに、検索された推薦対象音楽データが環境状況(天候)とユーザの感性情報に合致しているかどうかの評価を行った。それぞれの環境に対する相関量によるランキング上位20件において各楽曲が感性と一致しているかどうかを正解2点、準正解1点、不正解0点の三段階で被験者に評価させ、点数を百分率に直したものを適合率とする。結果を表4、表5、図6、図7に示す。図6および図7は、実際の楽曲抽出の結果の一部(上位10件)である。

表4、表5、図6、図7から、全体として、4つの計算方法のうち、どの計算方式から生成された感性・環境状況ベクトルによっても、5割以上の適合率を見せていることが分かる。なかでもとりわけ最大値を採用した場合では、平均して7割前後の精度を保っている。これは恐らく、最大値によって生成された感性・環境状況ベクトルが、他の方式では抽出しにくい微妙な感性情報を損なうことなく抽出していることにより、感性・音楽空間へのマッピングが成功し、楽曲推薦に反映することができているものと考えられる。このような微妙な感性情報を省略してしまうことなく抽出できるのは最大値を採用した場合のみであるため、音楽・感性マトリクス生成の際のユーザの感性情報の抽出方式として適しており、本実験によってそれぞれの印象語において最大値を採用することの有効性を示すことができる。

### 4.3. 考察

全体として、実験1における環境状況判定の結果は良好であり、入力画像を天候や周辺環境などの比較的オブジェクトに左右されない風景画像に限定すれば、色彩情報だけでも十分マッチングが取れることを表している。誤判定については、サンプル画像のデータ量を増やし、例外的な画像を排除することで、十分に精度の向上が可能と考えられる。

また、実験2における感性情報を媒介とした環境状況と音楽データのマッピングの結果は非常に良好であり、風景という人が認識しやすい感性特徴を利用し、天候や周辺環境などの風景画像に合致するサンプル楽曲を選択させることで、感性・環境状況マトリクスの自動生成が可能であることが分かった。これは、色彩を特徴とする風景画像データが、印象語を特徴とする音楽データに変換されることにより、環境状況が音楽感性の空間への写像が可能なデータとして変換されたとも考えることが出来る。使用するサンプル楽曲のジャンルを多岐にわたらせることで、より有効な感性・環境状況マトリクスの自動生成と楽曲推薦を行うことができると考えられる。

### 5. 結論と今後の展望

本稿では、ユーザの入力した風景画像の色彩情報から環境状況コンテキストを判定し、同時に、ユーザの選択したサンプル楽曲からユーザの嗜好・感性を抽出し、音楽感性属性を媒介とした環境状況コンテキストと音楽感性データの相関量計量により、コンテキストに相関が高く、個人の感性に合致した音楽データを推薦・提示するシステムの実現方式について示した。ま

た、空の画像を実風景として、それに相関の高い環境状況コンテキスト(ここでは天候)を判定し、その天候に対応する感性データの抽出を行い、その感性データと相関の高い音楽データの自動抽出を行う実験を行うことで、本方式の実現可能性を明らかにした。

今後の方針としては、実現方式の改良および実装の段階で、より細かい環境状況と感性の関連をセンシングするため、今回7語であった環境語を増やした状態での実験と並行して、前後の環境語の組み合わせによるストーリーの抽出といった視点からも研究を進めていく。また、今回の実験では、単一のアーティストでの楽曲推薦を行ったが、より実際の利用環境に近づけるべく、複数のジャンルの異なるアーティストの楽曲を使用した定量的実験を行い、本方式の有用性を示したいと考える。

本方式を発展させ、最終的に携帯電話に実装することで、携帯カメラで撮影した画像がそのまま、その時々ユーザの感性や感情の変化に応じたBGMに変換されるシステムを実現し、回線の高速化に伴う大量の楽曲配信に対応した実用型ユーザインタフェースとして展開することを考えている。

## 文 献

- [1] 土方嘉徳, “情報推薦・情報フィルタリングのためのユーザプロファイリング技術”, 人工知能学会, Vol.19, No.3, pp.365-372, 2004.
- [2] 小原恭介, 山田剛一, 細川博之, 中川裕志, “Bloggerの嗜好を利用した協調フィルタリングによるWeb情報推薦システム”, JSAI (人工知能学会), 2C2-02, 2005.
- [3] M. Anderson, M. Ball, H. Boley, S. Greene, N. Howse, D. Lemire and S. McGrath, “RACOFI: A Rule-Appling Collaborative Filtering System.” Proceedings of Workshop on Collaboration Agents: Autonomous Agents for Collaborative Environments, Oct. 2003.
- [4] 梶克彦, 平田圭二, 長尾確, “状況と嗜好に関するアンテーションに基づくオンライン楽曲推薦システム,” 2004-MUS-58, pp.33-38, 2004.
- [5] 竹川和毅, 土方嘉徳, 西田正吾, “内容に基づく音楽探索・推薦方式の実装,” 電子情報通信学会第18回データ工学ワークショップ(DEWS2007), 2007.
- [6] S. Pauws and B. Eggen, “Realization and user evaluation of an automatic playlist generator,” Journal of New Music Research, Volume 32, Number 2, pp. 179-192(14), June 2003.
- [7] 黒瀬崇弘, 梶川嘉延, 野村康雄, “感性情報を用いた楽曲推薦システム”, 第14回データ工学ワークショップ(DEWS2002), 8-P-6, 2003.
- [8] 小林重順, カラーイメージスケール改定版, 講談社, 2001.
- [9] 板橋美子, 佐々木史織, 吉田尚史, 清水 康, 時代・文化に応じた色・印象知識による文化財画像メタデータ自動抽出システムの実現方式, 日本データベース学会 Letters Vol.5, No.4, pp.21-24, March 2007.
- [10] Google Image Search: <http://images.google.co.jp/>
- [11] Flickr: <http://www.flickr.com/>
- [12] Kate Hevner, “Experimental Studies of the Elements of Expression in Music,” American Journal of Psychology, 48, pp. 246-268, 1936.
- [13] Kitagawa T. and Kiyoki Y. (2001) “Fundamental Framework for Media Data Retrieval Systems Using Media-lexico Transformation Operator - In the Case of Musical MIDI Data,” Information Modeling and Knowledge Bases (IOS Press), Vol. 12, pp. 316-326.